

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 22, 2005, 03:09:29 ; Search time 5708 Seconds  
(without alignments)  
10441.471 Million cell updates/sec

Title: US-10-690-564-1

Perfect score: 1230  
Sequence: 1 ccacgcgcctccgaacactca.....aaaaaaaaaaaaaaaaaaaa 1230

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl:\*  
1: gb\_ba:\*  
2: gb\_hcg:\*  
3: gb\_in:\*  
4: gb\_cm:\*  
5: gb\_ov:\*  
6: gb\_pac:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sbs:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vi:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1230	100.0	1230	8	AF083913	AF083913 Arabidops
2	1197.4	97.3	1208	8	AY072347	AY072347 Arabidops
3	1162.4	94.5	1192	8	AY086570	AY086570 Arabidops
4	1139	92.6	1139	8	BT003359	BT003359 Arabidops
5	1124.2	91.4	1206	8	ATANNEX	X99224 A.thaliana
6	1120	91.1	1159	8	ATU28415	U28415 Arabidopsis
7	1016.2	82.6	81513	8	FLAD7	AC021198 Sequence
8	954	77.6	954	6	CQ086050	Sequence
9	954	77.6	954	8	AF332435	AF332435 Arabidops
10	790.8	64.3	954	8	AY356355	AY356355 Brassica
11	497.8	40.5	1141	8	GHU89609	U89609 Gossypium h
12	488.6	39.7	948	8	GHU73746	GHU73746
13	478.6	38.9	1112	8	AF006197	AF006197 Lavatera
14	471.8	38.4	2588	8	BD236006	BD236006 Materials
15	468.6	38.1	1293	6	BD236034	BD236034 Materials
16	468.6	38.1	1293	6	AR231102	AR231102 Sequence
17	461.2	37.5	1180	8	CAANN	X93308 C.annuum mr
18	458	37.2	1420	8	BT012975	BT012975 Lycopersi
19	457.2	37.2	1160	8	AY085713	AY085713 Arabidops

20	455.6	37.0	1137	8	AF083914	AF083914 Arabidops
21	455.6	37.0	1230	8	AY070400	AY070400 Arabidops
22	453.2	36.8	954	6	CQ086274	CQ086274 Sequence
23	453.2	36.8	985	8	AY096577	AY096577 Arabidops
24	451	36.7	1080	8	CAN130956	AJ130956 Capsicum
25	450.6	36.6	1086	8	AY351650	AY351650 Gossypium
26	449.6	36.6	1157	6	AX927134	AX927134 Sequence
27	449.6	36.6	1197	8	AF113545	AF113545 Nicotiana
28	449.2	36.5	945	8	STU401032	AJ401032 Solanum t
29	447.6	36.4	1160	8	AF079232	AF079232 Lycopersi
30	445.4	36.2	1196	8	NTANNB511	Y14972 Nicotiana t
31	444.8	36.2	1182	8	MSANN	X74947 M.sativa mr
32	436.2	35.5	1172	8	NTANNB671	Y14973 Nicotiana t
33	430.4	35.0	1005	8	AY014799	AY014799 Arabidops
34	427.2	34.7	1045	8	AF079231	AF079231 Lycopersi
35	426.2	34.7	1146	8	GHU73747	U73747 Gossypium h
36	393.8	32.0	1003	8	AY014798	AY014798 Arabidops
37	393.8	32.0	1120	8	AK175572	AK175572 Arabidops
38	393.8	32.0	1120	8	AK175641	AK175641 Arabidops
39	393.8	32.0	1120	8	AK175892	AK175892 Arabidops
40	389	31.6	472	11	BX548120	BX548120 Arabidops
41	364.2	29.6	789	6	BD236004	BD236004 Materials
42	364.2	29.6	789	6	BD236037	BD236037 Materials
43	364.2	29.6	789	6	AR231105	AR231105 Sequence
44	346.6	28.2	1277	8	ZMANP33	X98244 Z.mays mRNA
45	344	28.0	1223	8	AK101787	AK101787 Oryza sat

#### ALIGNMENTS

RESULT 1  
AF083913  
LOCUS AF083913 1230 bp mRNA linear PIN 02-JUN-1999  
DEFINITION Arabidopsis thaliana annexin (AnnA1) mRNA, complete cds.  
ACCESSION AF083913  
VERSION AF083913.1 GI:4959105  
KEYWORDS  
SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopses.  
1 (bases 1 to 1230)  
Clark,G.B. and Roux,S.J.  
Isolation and Characterization of Two Different Arabidopsis Annexin  
cDNAs (Accession Nos. AF083913 and AF083914) (PCR 99-065)  
JOURNAL  
TITLE  
AUTHORS  
REFERENCE  
CDS  
gene  
1..1230  
/gene="AnnA1"  
34..987  
/gene="AnnA1"  
/note="Calcium-binding protein"  
/codon\_start=1  
/product="annexin"  
/protein\_id="AAD34236.1"  
/db\_xref="GI:4959106"  
/translation="MATLKVSDSPAPSDDAEQLTAFEGMCTNEDLIITSIARSAE  
ORKVIRQAVHYETGSDILKTLDKESNDPERAILMTLEPGERDLINAEATKWTSS  
NOVLMEVATRTSTOLHARQAYHARVYKSLSEEDVAHHTTGGFRLLVLSVSYREG  
DEVNMTLAKQSAKLVHEKIKQHYNDEIVIRLSTRSAQAINATFNRODDHGEILK  
SLEEGDDDKFLALRSTIGCLTRELTFVYVLRANKTGTDEGALRIYVTTAEID

## ORIGIN

LNVIGBYORNSIPLEKAIYKDTRGDYKMLVALLGEBDA"

Query Match 100.0%; Score 1230; DB 8; Length 1230;  
 Best Local Similarity 100.0%; Pred. No. 8, 1e-301;  
 Matches 1230; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY      1 CCACGGCTCCGAAACATAAAATAGAGAAAAATGGCGACTTTAAGGTTTCTGATTC 60
DB      1 CCACGGCTCCGAAACATAAAATAGAGAAAAATGGCGACTTTAAGGTTTCTGATTC 60
QY      61 GTTCCCTGCTCTTCTGATGATGCTGAGCAATTGAGACCGCTTTTGAAGATGGGTACG 120
DB      61 GTTCCCTGCTCTTCTGATGATGCTGAGCAATTGAGACCGCTTTTGAAGATGGGTACG 120
QY      121 AACGAGACTTGTATCATATCAATCTTGGCTCACAGAAAGTGTGAACAGAGAAAGTCATC 180
DB      121 AACGAGACTTGTATCATATCAATCTTGGCTCACAGAAAGTGTGAACAGAGAAAGTCATC 180
QY      181 AGGCAAGCATACCAAGCAAACTACGCGAAGACCTTCTCAAGACTTGTACAAGAGCTC 240
DB      181 AGGCAAGCATACCAAGCAAACTACGCGAAGACCTTCTCAAGACTTGTACAAGAGCTC 240
QY      241 TCTAAGATTTTGAAGAGACTATCTTGTGTGAGACTTCTTGAACCCGCTGAGCGTGATGCT 300
DB      241 TCTAAGATTTTGAAGAGACTATCTTGTGTGAGACTTCTTGAACCCGCTGAGCGTGATGCT 300
QY      301 TTATTGGCTAATGAACTCAAAAAGATGSACTTCAAGAACCAAGTTCTTATGAAAGTT 360
DB      301 TTATTGGCTAATGAACTCAAAAAGATGSACTTCAAGAACCAAGTTCTTATGAAAGTT 360
QY      361 GCTTGACAAGAGCATCAACGAGCTGCTTCAACGCTGAGCAAGCTTACCATGCTGCTAC 420
DB      361 GCTTGACAAGAGCATCAACGAGCTGCTTCAACGCTGAGCAAGCTTACCATGCTGCTAC 420
QY      421 AAGAACTCTCTTGAAGAGAGCTTGTCTCACACACTACCGGTGACTTCGAAAAGCTTTTG 480
DB      421 AAGAACTCTCTTGAAGAGAGCTTGTCTCACACACTACCGGTGACTTCGAAAAGCTTTTG 480
QY      481 GTTTCCTCTGTTACCTCATACAGGTGTCAGAAAGATGAAGTGAACATGCTGCTAAG 540
DB      481 GTTTCCTCTGTTACCTCATACAGGTGTCAGAAAGATGAAGTGAACATGCTGCTAAG 540
QY      541 CAAGAAAGCTAAGCTGTGCTCATGAGAAATCAAGACCAAGCACTACATGATGAGATGTT 600
DB      541 CAAGAAAGCTAAGCTGTGCTCATGAGAAATCAAGACCAAGCACTACATGATGAGATGTT 600
QY      601 ATTAGAATTTGTCCCAAGAAAGCAAGCTCAGATCAATGCTACTTTTAAACGTTAACCA 660
DB      601 ATTAGAATTTGTCCCAAGAAAGCAAGCTCAGATCAATGCTACTTTTAAACGTTAACCA 660
QY      661 GATGATCATGCGAGGAATTTCTCAAGAGCTTTGAGAGAGATGATGATGACAAGTTC 720
DB      661 GATGATCATGCGAGGAATTTCTCAAGAGCTTTGAGAGAGATGATGATGACAAGTTC 720
QY      721 CTTGCACTTTTGAAGTCAACCATTCAGTCTTGAACAAGCAAGACTTTTACTTTGTGAT 780
DB      721 CTTGCACTTTTGAAGTCAACCATTCAGTCTTGAACAAGCAAGACTTTTACTTTGTGAT 780
QY      781 GTTCTTCCTTGAAGCAATCAAAAATCTGAACTGATGAAAGGCACTCACTGAAATTTGG 840
DB      781 GTTCTTCCTTGAAGCAATCAAAAATCTGAACTGATGAAAGGCACTCACTGAAATTTGG 840
QY      841 ACCACAAGAGCTGAGATTGACTTGAAGTCAATGAGAGAGATACCAAGCGCAGAAACGC 900
DB      841 ACCACAAGAGCTGAGATTGACTTGAAGTCAATGAGAGAGATACCAAGCGCAGAAACGC 900
QY      901 ATTCTTTTGAAGAAAGCTATTACCAAGACACTCTGTGAGATTACGAGAAAGATGCTGTC 960
DB      901 ATTCTTTTGAAGAAAGCTATTACCAAGACACTCTGTGAGATTACGAGAAAGATGCTGTC 960
QY      961 GCACCTTCTGCGTGAAGATGATGCTTAATCAATCAATCTCTCCACAGAGAAACATAAGCTGC 1020
DB      961 GCACCTTCTGCGTGAAGATGATGCTTAATCAATCAATCTCTCTCCACAGAGAAACATAAGCTGC 1020

```

```

DB      961 GCACCTTCTGCGTGAAGATGATGCTTAATCAATCAATCTCTCCACAGAGAAACATAAGCTGC 1020
QY      1021 TCTACAGCTTCTGTTATCTTATCTCCCTCTCTCTCTTGAAGATTTCAATCGTT 1080
DB      1021 TCTACAGCTTCTGTTATCTTATCTCCCTCTCTCTCTTGAAGATTTCAATCGTT 1080
QY      1081 TGATTTTGTCTTACAAAACCTTGTGTTCTGTGTGTGTTTGAAGTTCCATAATTA 1140
DB      1081 TGATTTTGTCTTACAAAACCTTGTGTTCTGTGTGTGTTTGAAGTTCCATAATTA 1140
QY      1141 TGCAAAAGAGAGACAGAGAGACCAAGTGTGCTCTTAACTATATATATGAAGG 1200
DB      1141 TGCAAAAGAGAGACAGAGAGACCAAGTGTGCTCTTAACTATATATATGAAGG 1200
QY      1201 CATGGCCCTAAACAAAAA 1230
DB      1201 CATGGCCCTAAACAAAAA 1230

```

RESULT 2  
 AY072347  
 LOCUS  
 DEFINITION  
 Arabidopsis thaliana Ca2+-dependent membrane-binding protein annexin (At1g35720) mRNA, complete cds.  
 ACCESSION  
 AY072347  
 VERSION  
 AY072347.1 GI:18252242  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Arabidopsis thaliana (chale crese)  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE  
 1 (bases 1 to 1208)  
 Nguyen M., Karlin-Neumann G., Southwick A., Lam B., Miranda M., Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H., Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J., Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H., Sakurai T., Saitou M., Seki M., Shim P., Yamada K., Shinzaki K., Ecker J., Theologis A. and Davis R.W.

TITLE  
 JOURNAL  
 Submitted (02-JAN-2002) DNA Sequencing and Technology Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA

COMMENT  
 e-mail for correspondence: arab@sequence.stanford.edu

RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN Arabidopsis Full-length cDNA') : Seki M., Narusaka M., Ishida J., Saitou M., Kamiya A., Sakurai T., Carninci P., Kawai J., Hayashizaki Y. and Shinzaki K.

The Salk, Stanford, RGC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Nguyen M., Southwick A., Karlin-Neumann G., Lam B., Miranda M., Palm C.J., Bowser L., Jones T., Banh J., Chen H., Cheuk R., Chung M.K., Kim C., Lin J., Liu S.X., Pham P.K., Sakano H., Shim P., Yamada K., Ecker J., Theologis A. and Davis R.W.

Nguyen M. (SSP/Stanford) and Seki M. (RIKEN GSC) contributed equally to this work. Shinzaki K. (RIKEN GSC) and Davis R.W. (SSP/Stanford) contributed equally to this work as PIs.

FEATURES  
 Location/Qualifiers

source

1..1208

/organism="Arabidopsis thaliana"

/mol\_type="mRNA"

/db\_xref="taxon:3702"

/chromosome="1"

/clone="RAFL06-11-P04"

/ecotype="Columbia"

/note="This clone is in pBluescript"

1..1208

/gene="At1g35720"

33..986

gene

CDS

```

/gene="AL1935720"
/codon_start=1
/product="Ca2+-dependent membrane-binding protein annexin"
/protein_id="AL1954.1"
/db_xref="GI:18252243"
/translation="MATLKVSDVPAPSDAEQRTAFEGMGTEDELLIISILHRSAB
ORVIAOAHETGYEDLTKLDEKLSNDERAILMTHEGEDEDLALAEATRTMSS
NOVIMEVACTRTSTOLLHAROAHYARKSSLEDEVAHHTGTGPRKLLVSLNYSYREG
DEVNMTLAKOEAKLVHEKIKDRAYNDEEDVTRILSTRSKQINATFRYDDHEEELK
SLEEGDDKFLALRSTIQCLTRPELVFVDVLRSAINTGTGALTRIVTTRAEID
LKVIIEYORRNSIPLEKAIYKDTGRDYKMLVALLGDDA"

```

## ORIGIN

```

Query Match      97.3%; Score 1197.4; DB 8; Length 1208;
Best Local Similarity 99.9%; Pred. No. 1,5e-292;
Matches 1198; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 10 CGAACAACATAAAGTGAAGAAAAATGCGCACTTAAGGTTCTGATTCCTGCTGCT 69
    |||
Db 9 CAAAACCTRAAAGTGAAGAAAAATGCGCACTTAAGGTTCTGATTCCTGCTGCT 68

QY 70 CCTTGTGATGATGCTGAGCAATTGAAACCGCTTTTGAAAGATGGGCTACGAAGA 129
    |||
Db 69 CCTTGTGATGATGCTGAGCAATTGAAACCGCTTTTGAAAGATGGGCTACGAAGA 128

QY 130 TTGATCATATCAATCTTTGGCTCAAGAGTCTGAAACAGAGAAAGTCATCAGGCAAGCA 189
    |||
Db 129 TTGATCATATCAATCTTTGGCTCAAGAGTCTGAAACAGAGAAAGTCATCAGGCAAGCA 188

QY 190 TACCAAGAAACCTACGCGCGAAGACCTTCTCAAGACTCTTGACAAGAGCTCTCAAGCAT 249
    |||
Db 189 TACCAAGAAACCTACGCGCGAAGACCTTCTCAAGACTCTTGACAAGAGCTCTCAAGCAT 248

QY 250 TTGAGAGAGCTATCTTGTGTGAGACTCTTGAAACCGGTGAGCGGTATGCTTTATTTGGCT 309
    |||
Db 249 TTGAGAGAGCTATCTTGTGTGAGACTCTTGAAACCGGTGAGCGGTATGCTTTATTTGGCT 308

QY 310 AATGAAGCTACAAAAAGATGAGACTTCAAGAACCAAGTTTATGGAAGTCTGCTGCA 369
    |||
Db 309 AATGAAGCTACAAAAAGATGAGACTTCAAGAACCAAGTTTATGGAAGTCTGCTGCA 368

QY 370 AGGACATCAACGCGAGCTTCAACGCTAAGGCAAGCTTACCATGCTGCTACAGAAAGTCT 429
    |||
Db 369 AGGACATCAACGCGAGCTTCAACGCTAAGGCAAGCTTACCATGCTGCTACAGAAAGTCT 428

QY 430 CTTGAAGAGAGCGTTGCTCAACACACTACCGGTGACTTCAAGAAAGCTTTTGTCTCTT 489
    |||
Db 429 CTTGAAGAGAGCGTTGCTCAACACACTACCGGTGACTTCAAGAAAGCTTTTGTCTCTT 488

QY 490 GTTACTCTATCAAGGTAACGAGAGATGAAGTAACATGACATGGCTTAAGCAAGAAAGCT 549
    |||
Db 489 GTTACTCTATCAAGGTAACGAGAGATGAAGTAACATGACATGGCTTAAGCAAGAAAGCT 548

QY 550 AAGCTGTGCATAGAAAAATCAAGACAAGCACTACATAGATGAGATGTTATTAGAATC 609
    |||
Db 549 AAGCTGTGCATAGAAAAATCAAGACAAGCACTACATAGATGAGATGTTATTAGAATC 608

QY 610 TTGTCCACAAGAGCAAGGCTCAGATCAATGCTATCTTTTAACCGTTACCAAGATGATCAT 669
    |||
Db 609 TTGTCCACAAGAGCAAGGCTCAGATCAATGCTATCTTTTAACCGTTACCAAGATGATCAT 668

QY 670 GGGGAGAAATTTCTCAAGAGCTTTGAGAGAAAGATGATGATCAAGATTCCTTGAACCTT 729
    |||
Db 669 GGGGAGAAATTTCTCAAGAGCTTTGAGAGAAAGATGATGATCAAGATTCCTTGAACCTT 728

QY 730 TTGAGGTCAACATTAAGTCTTGAACAAGCAAGAGCTTACTTTCGATGTTCTTCGT 789
    |||
Db 729 TTGAGGTCAACATTAAGTCTTGAACAAGCAAGAGCTTACTTTCGATGTTCTTCGT 788

QY 790 TCAGCAATCAACAAAACCTGAACTGATGAGAGCACTCACTAAGATTGACCAACAAG 849
    |||
Db 789 TCAGCAATCAACAAAACCTGAACTGATGAGAGCACTCACTAAGATTGACCAACAAG 848

```

```

QY 850 GCTGAGATTGACTTGAAGGTCATTTGAGAGAGATCACAGCGAGGAACGATTCCTTG 909
    |||
Db 849 GCTGAGATTGACTTGAAGGTCATTTGAGAGAGATCACAGCGAGGAACGATTCCTTG 908

QY 910 GAGAAAGCTATTATCCAAAGACACCTGAGAGATTAAGAGAAAGTCTGCTGCACTTCTC 969
    |||
Db 909 GAGAAAGCTATTATCCAAAGACACCTGAGAGATTAAGAGAAAGTCTGCTGCACTTCTC 968

QY 970 GGTGAAGATGATGCTTAATCAATCAATCTCCACAGAGAAACATAAGCTGCTTACAGCT 1029
    |||
Db 969 GGTGAAGATGATGCTTAATCAATCAATCTCCACAGAGAAACATAAGCTGCTTACAGCT 1028

QY 1030 TCTGTTATCTTATATCCCTCTCTCTCTCTCTTGTGATGAGTTCAATCGTTGATTTGT 1089
    |||
Db 1029 TCTGTTATCTTATATCCCTCTCTCTCTCTCTTGTGATGAGTTCAATCGTTGATTTGT 1088

QY 1090 TTCTACAAAACCTTGTGTTGTTCTGTTGTTGTTGTTGATTCTTAATATGCAAAAG 1149
    |||
Db 1089 TTCTACAAAACCTTGTGTTGTTCTGTTGTTGTTGATTCTTAATATGCAAAAG 1148

QY 1150 GAGAGACAGAGAAACCAAGTGTGCTCTTAAGTATATATATGAGAGCAATTGGCC 1208
    |||
Db 1149 GAGAGACAGAGAAACCAAGTGTGCTCTTAAGTATATATATGAGAGCAATTGGCC 1207

```

```

RESULT 3
AY086570      1192 bp  mRNA  linear  PLN 14-APR-2003
LOCUS        Arabidopsis thaliana clone 25846 mRNA, complete sequence.
DEFINITION   AY086570
ACCESSION    AY086570.1  GI:21405280
VERSION      AY086570.1  GI:21405280
KEYWORDS     FLI CDNA.
SOURCE       Arabidopsis thaliana (thale cress)
ORGANISM     Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosids; eustosids II; Brassicales; Brassicaceae; Arabidopsids.
1 (bases 1 to 1192)
Haas, B. J., Volkovskiy, N., Town, C. D., Troupkhan, M., Alexandrov, N.,
Feldmann, K. A., Flavell, R. B., White, O. and Salzberg, S. L.
Full-length messenger RNA sequences greatly improve genome
annotation
Genome Biol. 3 (6), RESEARCH0029 (2002)
22088475
JOURNAL
MEDLINE
PUBMED
2 (bases 1 to 1192)
12093376
REFERENCE
Broyer, V., Troupkhan, M., Alexandrov, N., Lu, Y.-P., Flavell, R. and
Feldmann, K.
Full-length cDNA from Arabidopsis thaliana
Unpublished
3 (bases 1 to 1192)
Broyer, V., Troupkhan, M., Alexandrov, N., Lu, Y.-P., Flavell, R. and
Feldmann, K.
Direct Submision
Submitted (11-MAR-2002) Ceres, Inc. 3007 Malibu Canyon Road,
Malibu, CA 90265, USA
This clone sequence is one of 5,000 Ceres full-length cDNAs made
available to TIGR and Genbank. The following quality assessment of
this set was done by comparison with known proteins: two percent
of the clones are estimated to be 5'-truncated; less than one percent
are 3'-truncated; approximately two percent represent alternative
splice variants, including unspliced introns and spliced exons; one
percent may contain premature stop codons; five percent may have
frame shifts in a coding region. A sequence is considered to be
5'-truncated if it lacks the translation initiation start (ATG). A
sequence is considered to be 3'-truncated if it lacks the
C-terminal end of the encoded protein. Please note that these cDNA
sequences are derived from the Ws or Laer ecotypes and therefore
may contain polymorphisms when compared to sequences from Col-0.
Geneset carried out the library production and sequencing of the
full-length clones. Ceres, Inc. carried out the clustering of the
5' sequences, selection of clones, and sequence assembly.
Location/Qualifiers

```

## FEATURES

source	1. .1192
	/organism="Arabidopsis thaliana"
	/mol_type="mRNA"
	/db_xref="taxon:3702"
	/clone="25846"
CDS	36. .989
	/codon_start=1
	/product="Ca2+-dependent membrane-binding protein annexin"
	/protein_id="AA063633.1"
	/db_xref="GI:21554612"
	/translation="MATTAKVSDSPAPSDAQLRTAREGKNTEDLIISIAHRSAS QKRVLRQAYHETGDELTKLDEKLSNDERAIIILMTLEPERDALLINAEKRWTS NOVMEVACTRTSTOLHARQAYHARKYKSELDVAHHTTGPRLVSLVSYVEG DEVNMTLAKQEAHLVHEKIKDHYNDEVDYRIISTRSKQINATYRQVDDGEILK SLEGGDDDKFLALRSTIGCTREPELVVDLRSAINKTGDEGALRIVYTRAID LKVIEEYRRRNSIPLEKAITDTRGDYERKMLVALLGEDA"
ORIGIN	
Query Match	94.5%; Score 1162.4; DB 8; Length 1192;
Best Local Similarity	99.1%; Pred. No. 1.1e-283;
Matches 1169; Conservative	0; Mismatches 11; Indels 0; Gaps 0;
QY	10 CGAAACACTAAAGTGAAGAAAAATGGCGACTTTAAGCTTTCTGATTCCTGCT 69
DB	12 CAAACACTAAAGTGAAGAAAAATGGCGACTTTAAGCTTTCTGATTCCTGCT 71
QY	70 CCTTCGATGATGCTGAGCAATTTGAGAACCGCTTTTGAAGGATGGGGTACAGAGCAG 129
DB	72 CCTTCGATGATGCTGAGCAATTTGAGAACCGCTTTTGAAGGATGGGGTACAGAGCAG 131
QY	130 TTGATCATATCAATCTTGGCTCAGAGAGTGTGAACAGAGGAAGTCAATCAGGCAAGCA 189
DB	132 TTGATCATATCAATCTTGGCTCAGAGAGTGTGAACAGAGGAAGTCAAGGCAAGCA 191
QY	190 TACCAAGAACTTACGGCGCAAGACCTTCTCAAGACTCTTGACAAGAGCTCTCTAACAT 249
DB	192 TACCAAGAACTTACGGCGCAAGACCTTCTCAAGACTCTTGACAAGAGCTCTCTAACAGC 251
QY	250 TTTCGAGAGAGCTATCTTGTGTGAGACTCTTGAACCGGAGCGAGTATCTTAAATGGCT 309
DB	252 TTTCGAGAGAGCTATCTTGTGTGAGACTCTTGAACCGGAGCGAGTATCTTAAATGGCT 311
QY	310 AATGAAGCTACAAAAGATGACTTCAAGCAACCAAGTTCTTAAGAAATTGCTTGACA 369
DB	312 AATGAAGCTACAAAAGATGACTTCAAGCAACCAAGTTCTTAAGAAATTGCTTGACA 371
QY	370 AGGACATCAACGCGAGCTGCTTACCGTGAAGCAAGCTTACATGCTGCTCAAGAAAGCT 429
DB	372 AGGACATCAACGCGAGCTGCTTACCGGAGCAAGCTTACATGCTGCTCAAGAAAGCT 431
QY	430 CTGGAAGAGAGCTTGCTCAGCACTACCGGTGACTTCAAGAAAGCTTTGGTTCTCTT 489
DB	432 CTGGAAGAGAGCTTGCTCAGCACTACCGGTGACTTCAAGAAAGCTTTGGTTCTCTT 491
QY	490 GTTACCTCATACAGGTACGAGAGAGATGAAGTGAACATGACATTTGGCTAAGCAAGAGCT 549
DB	492 GTTACCTCATACAGGTACGAGAGAGATGAAGTGAACATGACATTTGGCTAAGCAAGAGCT 551
QY	550 AAGCGTGTCAAGAAAAATCAAGACAAGCACTACATGATGAGAGATTTTATGATC 609
DB	552 AAGCGTGTCAAGAAAAATCAAGACAAGCACTACATGATGAGAGATTTTATGATC 611
QY	610 TTGTCACACAGAGCAAGCTCAGATCAATGTCACTTTTAAACGTTTACCAATGATGAT 669
DB	612 TTGTCACACAGAGCAAGCTCAGATCAATGTCACTTTTAAACGTTTACCAATGATGAT 671
QY	670 GCGGAGAAATTTCTCAAGAGCTTTGAAGAGAGATGATGATGACAAAGTTCTTGCACAT 729
DB	672 GCGGAGAAATTTCTCAAGAGCTTTGAAGAGAGATGATGATGACAAAGTTCTTGCACAT 731
QY	730 TTGAGCTCAACGATTCAGTCTTGACACAGACAGAGCTTTACTTTGTGATGTTCTTGCT 789
DB	732 TTGAGCTCAACGATTCAGTCTTGACACAGACAGAGCTTTACTTTGTGATGTTCTTGCT 791

QY	790 TCAGCAATCAACAAACTGAACTGATGAGAGCACTCACTAGATTTGTGACACAAGA 849
DB	792 TCAGCAATCAACAAACTGAACTGATGAGAGCACTCACTAGATTTGTGACACAAGA 851
QY	850 GCTGAGATTGACTTGAAGTCAATTGAGAGAGATACAGCGCAGGACAGACATTCCTTTG 909
DB	852 GCTGAGATTGACTTGAAGTCAATTGAGAGAGATACAGCGCAGGACAGACATTCCTTTG 911
QY	910 GAGAAAGCTATTAACCAAGACACTCTGTGAGATTAACGAAAGATTCGTGCACTTCTC 969
DB	912 GAGAAAGCTATTAACCAAGACACTCTGTGAGATTAACGAAAGATTCGTGCACTTCTC 971
QY	970 GGTGAAGATGATGCTTAATCAATCAATCTCCACAGAGAAACATTAAGCTGCTACAGCT 1029
DB	972 GGTGAAGATGATGCTTAATCAATCAATCTCCACAGAGAAACATTAAGCTGCTACAGCT 1031
QY	1030 TCTGTTATCTTATCTCCTCTCTCTCTCTTGTGATGAGATTCAAATCGTTGATTTGT 1089
DB	1032 TCTGTTATCTTATCTCCTCTCTCTCTCTTGTGATGAGATTCAAATCGTTGATTTGT 1091
QY	1090 TTCTACAAAACCTGTTGTTTCTGTTGTGTTGTTTGAAGTTCCATAATGCAAAAGA 1149
DB	1092 TTCTACAAAACCTGTTGTTTCTGTTGTGTTGTTTGAAGTTCCATAATGCAAAAGA 1151
QY	1150 GAGAGACAGAGAGAACACAGTGTGCTCTTAAGTTATATA 1189
DB	1152 GAGAGACAGAGAGAACACAGTGTGCTCTTAAGTTATATA 1191

RESULT 4	BT003359	1139 bp	mRNA	linear	PLN 29-JAN-2003
LOCUS	BT003359				
DEFINITION	Arabidopsis thaliana Ca2+-dependent membrane-binding protein annexin (At1g35720) mRNA, complete cds.				
ACCESSION	BT003359				
VERSION	BT003359.1	GI:28059005			
KEYWORDS	FLI CDNA.				
SOURCE	Arabidopsis thaliana (thale cress)				
ORGANISM	Arabidopsis thaliana				
REFERENCE	Southwick, A., Nguyen, M., Tripp, M., Palm, C.J., Jones, T., Wu, T., Deng, J.M., Hayashizaki, Y., Hsuan, V.W., Lee, J.M., Ishida, J., Kamliya, A., Kawai, J., Kim, C.J., Narusaka, M., Onodera, C.S., Quach, H.L., Sakurai, T., Satou, M., Seki, M., Shimizu, P., Tang, C.C., Torouni, M., Wong, C., Wu, H.C., Yamada, K., Yu, G., Yuan, S., Shinozaki, K., Ecker, J., Theologis, A. and Davis, R.W.				
AUTHORS	Direct Submission				
TITLE	Submitted (29-JAN-2003) DNA Sequencing and Technology Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA				
JOURNAL					
COMMENT	e-mail for correspondence: arab@sequence.stanford.edu				

RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN Arabidopsis Full-length cDNA') : Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamliya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.

The Sak, Stanford, PGSC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Nguyen, M., Tripp, M., Southwick, A., Palm, C.J., Jones, T., Wu, T., Chen, H., Cheuk, R., Chan, M.M., Chang, C.H., Dale, J.M., Deng, J.M., Hsuan, V.W., Lee, J.M., Kim, C.J., Quach, H.L., Onodera, C.S., Shimizu, P., Tang, C.C., Torouni, M., Wong, C., Wu, H.C., Yamada, K., Yu, G., Yuan, S., Ecker, J., Theologis, A. and Davis, R.W.

Nguyen, M., (SSP/Stanford) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Davis, R.W.





## ORIGIN

/db\_xref="UniProt/TREMBL:Q39001"  
 /translation="TLKSDSVPAESPDDAEOLRTAFEGMGTNEDLIISILAHRSAR  
 KYRQAYHETGSEDLAKTDKELSNPERRAILMTLEPERPAILANETKWTSSNO  
 VMEVACTRSTDLHARQAYHARKYKSLBEDVAHHTTGDPRFLVLSVTSRYEDE  
 VNMVTLAKQEKLVHEKIKDKNVDEDEVILSTISKAQINATFNRRQDDHGEELKSL  
 EBDGDDDKFALRSTIOCLTRPELVFVNLRSAINKTGTDEGALTRIVTTABEIDLK  
 VIGEEYQRNSIPLRAITKDNCGDYKMLVAFGEEDA"

Query Match 91.4%; Score 1124.2; DB 8; Length 1206;  
 Best Local Similarity 97.0%; Pred. No. 5,5e-274;  
 Matches 1157; Conservative 0; Mismatches 33; Indels 3; Gaps 1;

38 CGACTTTAAGGTTTCTGATTTCTGTCCTGCTCTGATGATGCTGAGCAATTGAGA 97  
 1 CGACTTTAAGGTTTCTGATTTCTGTCCTGCTCTGATGATGCTGAGCAATTGAGA 60

98 CCGCTTTTGAAGATGGGGTACGAACGAGGACTTGATCATATCATCTTGCTCAGAGA 157  
 61 CCGCTTTTGAAGATGGGGTACGAACGAGGACTTGATCATATCATCTTGCTCAGAGA 120

158 GTGCTGAACGAGAGAAAGTATCAGGCAAGCATACCAACCAACCTTACGGGAAAGCTTC 217  
 121 GTGCTGAACGAGAGAAAGTATCAGGCAAGCATACCAACCAACCTTACGGGAAAGCTTC 180

218 TCAAGACTCTTGAACAAGAGCTCTTACGATTTGAGAGAGCTATCTTGTTGGAGCTC 277  
 181 TCAAGACTCTTGAACAAGAGCTCTTACGATTTGAGAGAGCTATCTTGTTGGAGCTC 240

278 TTGAACCCGGTGAACGCTGATGCTTTATTTGGCTAATGAAAGTCAAAAAGATGACTTCA 337  
 241 TTGAACCCGGTGAACGCTGATGCTTTATTTGGCTAATGAAAGTCAAAAAGATGACTTCA 300

338 GCAACCAAGTTCTTTATGGAAGTTGCTTGCACAAGACATTAAGCAGCTTCAAGCTTA 397  
 301 GCAACCAAGTTCTTTATGGAAGTTGCTTGCACAAGACATTAAGCAGCTTCAAGCTTA 360

398 GCGAAGCTTCAAGTCCGCTACAAAGAGTCTCTGTAAGAGAGAGCTTGCACACACATTA 457  
 361 GCGAAGCTTCAAGTCCGCTACAAAGAGTCTCTGTAAGAGAGAGCTTGCACACACATTA 420

458 CCGGTGACTTCAGAAAGCTTTTGGTTCTCTTGTATCCTTATCAAGATGAGAGATG 517  
 421 CCGGTGACTTCAGAAAGCTTTTGGTTCTCTTGTATCCTTATCAAGATGAGAGATG 480

518 AATTGAACATGACATTTGCTTAACCAAGAAAGTAAAGCTGCTTGAAGAAATTAAGACA 577  
 481 AATTGAACATGACATTTGCTTAACCAAGAAAGTAAAGCTGCTTGAAGAAATTAAGACA 540

578 AGCACTACATGATGAGATGTTATTAGATCTTGTCCACAAGAAACCAAAAGCTCAGATCA 637  
 541 AGCACTACATGATGAGATGTTATTAGATCTTGTCCACAAGAAACCAAAAGCTCAGATCA 600

638 ATGCTACTTTTAAACCGTTACCAAGATGATCATGCGCAGGAAATTTCCAGAAGCTTGAAG 697  
 601 ATGCTACTTTTAAACCGTTACCAAGATGATCATGCGCAGGAAATTTCCAGAAGCTTGAAG 660

698 AAGAAGTATGATGACAAATTTCTTGCACATTTTGAAGTAAACCATTCAGTGTTCACAA 757  
 661 AAGAAGTATGATGACAAATTTCTTGCACATTTTGAAGTAAACCATTCAGTGTTCACAA 720

758 GACCAAGACTTTTAAACCGTTACCAAGATGATCATGCGCAGGAAATTTCCAGAAGCTTGAAG 817  
 721 GACCAAGACTTTTAAACCGTTACCAAGATGATCATGCGCAGGAAATTTCCAGAAGCTTGAAG 780

818 AAGAAGCACTCATTAGATTTGTGACCAACAAAGCTGAGATGACTTGAAGCTATTGAG 877  
 781 AAGAAGCACTCATTAGATTTGTGACCAACAAAGCTGAGATGACTTGAAGCTATTGAG 840

878 AGAAGTACCAAGCAAGCAATTCCTTTGAGAGAAAGTATTAACCAAGCAATTCCTG 937  
 841 AGAAGTACCAAGCAAGCAATTCCTTTGAGAGAAAGTATTAACCAAGCAATTCCTG 897

QY 938 GAGATTACGAGAGATGCTCGTCGCACTTTCGCTGAGAGATGATCTTAATCAATCAATC 997  
 DB 898 GAGATTACGAGAGATGCTCGTCGCACTTTCGCTGAGAGATGATCTTAATCAATCAATC 957

QY 998 CTCACAGAGAAACATTAAGCTGCTCTTACAGCTTCTGTTATCTCTTATCTCCCTCTCTC 1057  
 DB 958 CTCACAGAGAAACATTAAGCTGCTCTTACAGCTTCTGTTATCTCTTATCTCCCTCTCTC 1017

QY 1058 TCTTTGAGAGTTCAATTCGTTGATTTTGTCTTACAAAACCTTGTTGTTCTGTT 1117  
 DB 1018 TCTTTGAGAGTTCAATTCGTTGATTTTGTCTTACAAAACCTTGTTGTTCTGTT 1077

QY 1118 GTGTGTTTGAAGTCTTAATATGCAAAAGAGAGACAGAGAACAGTGTGCTC 1177  
 DB 1078 GTGTGTTTGAAGTCTTAATATGCAAAAGAGAGACAGAGAACAGTGTGCTT 1137

QY 1178 TTAAGTTATATATATGAGAGCACTTGGCTTAAATTAATTAATTAATTAATTAATTA 1230  
 DB 1138 TTAAGTTATATATATGAGAGCACTTGGCTTAAATTAATTAATTAATTAATTAATTA 1190

RESULT 6  
 ATU28415 1159 bp mRNA linear PIN 21-OCT-1996  
 LOCUS Arabidopsis thaliana annexin-like protein mRNA, complete cds.  
 DEFINITION U28415  
 ACCESSION U28415  
 VERSION U28415.1 GI:1621538  
 KEYWORDS

SOURCE Arabidopsis thaliana (thale cress)  
 ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE 1 (bases 1 to 1159)  
 Gidrol,X., Sabelli,P.A., Fern,Y.S. and Kush,A.K.  
 Annexin-like protein from Arabidopsis thaliana rescues delta oxyr  
 mutant of Escherichia coli from H2O2 stress  
 Proc. Natl. Acad. Sci. U.S.A. 93 (20), 11268-11273 (1996)  
 MEDLINE 97008169  
 PUBMED 8855345

REFERENCE 2 (bases 1 to 1159)  
 Kush,A.K.  
 Direct Submission  
 Submitted (06-JUN-1995) Anil K. Kush, Institute of Molecular & Cell  
 Biology, National University of Singapore, 10 Kent Ridge Crescent,  
 Singapore 0511, Republic of Singapore

FEATURES  
 source location/Qualifiers  
 1..1159  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /variety="ecotype Columbia"  
 /db\_xref="taxon:3702"

/clone="pOxy5"  
 /cissue\_type="whole seedling"  
 /clone\_lib="Lambda UnizapII cDNA library of A.K.Kush"  
 /dev\_stage="4 week-old"  
 6..959  
 /note="annexin-like protein"  
 /codon\_start=1  
 /protein\_id="AAC49472.1"  
 /db\_xref="GI:1621539"

CDS  
 /translation="MATLKSDSVPAESPDDAEOLRTAFEGMGTNEDLIISILAHRSAR  
 ORKVIROYHETGSEDLAKTDKELSNPERRAILMTLEPERPAILANETKWTSS  
 NOVMEVACTRSTDLHARQAYHARKYKSLBEDVAHHTTGDPRFLVLSVTSRYEYG  
 DEVMVTLAKQEKLVHEKIKDKNVDEDEVILSTISKAQINATFNRRQDDHGEELKSL  
 EBDGDDDKFALRSTIOCLTRPELVFVNLRSAINKTGTDEGALTRIVTTABEIDLK  
 VIGEEYQRNSIPLRAITKDNCGDYKMLVAFGEEDA"

## ORIGIN

Query Match 91.1%; Score 1120; DB 8; Length 1159;  
 Best Local Similarity 98.9%; Pred. No. 6,4e-273;  
 Matches 1149; Conservative 0; Mismatches 10; Indels 3; Gaps 2;

Oy	29	GAAAAATGGCGACTCTTAAGGTTTGTGATTCGTTCCGTCCTTCTGATGATGCTGACG	88
Db	1	GAAAAATGGCGACTCTTAAGGTTTGTGATTCGTTCCGTCCTTCTGATGATGCTGACG	60
Oy	89	AATTGAGAACCGCTTTTGAAGATGGGGTACGACGAGACTTGTATCATCAATCTTGG	148
Db	61	AATTGAGAACCGCTTTTGAAGATGGGGTACGACGAGACTTGTATCATCAATCTTGG	120
Oy	149	CTCACAGAAATGCTGACACAGGAAAGTCATCAGCAAGCATACACGAAACTTACGGCG	208
Db	121	CTCACAGAAATGCTGACACAGGAAAGTCATCAGCAAGCATACACGAAACTTACGGCA	180
Oy	209	AAGACCTTCTCAAGACTCTTTCGACAGAGAGCTCTCAAGATTGAGAGAGCTATCTTGT	268
Db	181	AAGACCTTCTCAAGACTCTTTCGACAGAGAGCTCTCAAGAGATTGAGAGAGCTATCTTGT	240
Oy	269	TGTGACTCTTGAACCCGGTGAAGCTGATGCTTTTGTGGCTAATGAACCTTCAAAAAGAT	328
Db	241	TGTGACTCTTGAACCCGGTGAAGCTGATGCTTTTGTGGCTAATGAACCTTCAAAAAGAT	300
Oy	329	GGACTTCAAGCAACCAAGTCTTATGAGAAATGTCCTTTCACAGACATCAACGACCTGC	388
Db	301	GGACTTCAAGCAACCAAGTCTTATGAGAAATGTCCTTTCACAGACATCAACGACCTGC	360
Oy	389	TTTACGCTAGCAAGCTTACCAATGTCGCTACAGAAATCTCTTGAAGAGAGCTTGTCTC	448
Db	361	TTTACGCTAGCAAGCTTACCAATGTCGCTACAGAAATCTCTTGAAGAGAGCTTGTCTC	420
Oy	449	ACCACATCACCGGTACTCTTCAAGAAAGCTTTGGTTCCTTGTTACCTATACAGGTACG	508
Db	421	ACCACATCACCGGTACTCTTCAAGAAAGCTTTGGTTCCTTGTTACCTATACAGGTACG	480
Oy	509	AAGGAGATGAAGTGAACATGACATTGGCTTAAGCAAGAACTAAGCTGTCATGAGAAA	568
Db	481	AAGGAGATGAAGTGAACATGACATTGGCTTAAGCAAGAACTAAGCTGTCATGAGAAA	540
Oy	569	TCAAGGACAGACATCAATGATGAGATGTTATTGAATCTTGTCCACAGAAAGCAAG	628
Db	541	TCAAGGACAGACATCAATGATGAGATGTTATTGAATCTTGTCCACAGAAAGCAAG	600
Oy	629	CTCAGATCAATGCTTACTTTTAAACGTTTACCAAGATGATCATGGGAGAAATTTCCAGA	688
Db	601	CTCAGATCAATGCTTACTTTTAAACGTTTACCAAGATGATCATGGGAGAAATTTCCAGA	660
Oy	689	GTCCTTGAGGAAGAGATGATGATGACAAATTCCTTTCACATTTTGAAGTCAACATTCAGT	748
Db	661	GTCCTTGAGGAAGAGATGATGATGACAAATTCCTTTCACATTTTGAAGTCAACATTCAGT	720
Oy	749	GCTTTCAGAACACAGAGCTTTTACTTTGTGATGTTCTTGTTCAGCAATCAACAAACTG	808
Db	721	GCTTTCAGAACACAGAGCTTTTACTTTGTGATGTTCTTGTTCAGCAATCAACAAACTG	780
Oy	809	GAACTGATGAAGAGGACCTCACTGAAATTTGACACACAGAGCTGAGTTGACTTGAAG	868
Db	781	GAACTGATGAAGAGGACCTCACTGAAATTTGACACACAGAGCTGAGTTGACTTGAAG	840
Oy	869	TCAATTGAGAGAGATACAGCGCAGGAACAGCATTCCTTTGAGAAAGCTATTACCAAG	928
Db	841	TCAATTGAGAGAGATACAGCGCAGGAACAGCATTCCTTTGAGAAAGCTATTACCAAG	900
Oy	929	ACACTCGTGAAGATTAAGAGAAATGCTCGTGCACATTCCTGCGTGAAGATGATCTTAAT	988
Db	901	ACACTCGTGAAGATTAAGAGAAATGCTCGTGCACATTCCTGCGTGAAGATGATCTTAAT	960
Oy	989	CAATCAATCCCTCCACAGAGAAACATTAAGCTGCTCAAGCTTCTGTATCTCTTATCTCC	1048
Db	961	CAATCAATCCCTCCACAGAGAAACATTAAGCTGCTCAAGCTTCTGTATCTCTTATCTCC	1019
Oy	1049	CTCTCTCTCTTTGATGATTTCAATCGTTTGAATTTGTTTCTACAAAAACCTTGTT	1108
Db	1020	CCCTCTCTCTTTGATGATTTCAATCGTTTGAATTTGTTTCTACAAAAACCTTGTT	1079
Oy	1109	GTTTCTGTGTGTGTTTGAAGTCTCTAATAATGCAAAAGAGAGACAGAGAAACGAG	1168

D<sub>b</sub> 1080 GTTCTGTGTCGTTTGACTCTCTAAATATGCAG--AGACAGCAGAGAACCG 1137  
Q<sub>y</sub> 1169 TGTGTCCTTAAGTATATAT |||||  
D<sub>b</sub> 1138 TGTCGTCCTTAAGTATATAT ||||| 1159

RESULT 7	
LOCUS	F14D7
DEFINITION	81513 bp DNA linear PLN 28-JUN-2006
ACCESSION	Sequence of BAC F14D7 from <i>Arabidopsis thaliana</i> chromosome 1, complete sequence.
VERSION	AC021198
KEYWORDS	AC021198.2 GI:6957696
SOURCE	HTG.
ORGANISM	<i>Arabidopsis thaliana</i> (thale cress) <i>Arabidopsis thaliana</i>

REFERENCE  
1 (bases 1 to 81513)  
Liu, S. X., Chan, A., Sakano, H., Yu, G., Lee, J. M., Lenz, C., Pham, P., et al., 2013. The genome of the Asian house mosquito, *Culex tritaeniorhynchus*. *Nature* 501: 521-525.

TITLE The sequence of BAC F14D7 from Arabidopsis thaliana chromosome 1

JOURNAL. Submitted (15-JAN-2000) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA  
REFERENCE 3 (bases 1 to 81513)

JOURNAL  
Submitted (09-FEB-2007) Plant Gene Expression Center, 800 Buchanan  
Street, Albany, CA 94710, USA  
4 (bases 1 to 81513)  
REFERENCE

COMMENT  
JOURNAL  
Submitted (28-JUN-2000) Plant Gene Expression Center, 800 Buchanan  
St., Albany, CA 94710, USA  
On Feb 11, 2000 this sequence version replaced gi:6693723.

The sequence is of BAC FlAD7 from *Arabidopsis thaliana* chromosome 1. The sequence does not represent the sequence of the entire insert of this clone. It is shorter by 594 bp because we submit only the unique sequence of the clone. However, in order to facilitate the joining of overlapping clones in the future for creation of larger contigs, we provide small overlaps (200 bp) between overlapping subcloned clones. The 5' end of this sequence overlaps by 200 bp to the 3' end of the sequence of the clone Fl504.

```

FEATURES
source
location/Qualifiers
1..81513
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/cultivar="Columbia"
/db_xref="taxon:3702"
/chromosome="1"
/clone="F14D7"
1553..4999
/gene="F14D7.1"
join(1553..4628,4713..4999)
CDS

```

```

/note="Contains similarity to receptor protein kinase-like
protein from Arabidopsis thaliana gb|A1.61513. It
contains a eukaryotic protein kinase domain Pf100069. EST
gb|A1997574 comes from this gene."
/codon_start=1

```

[illegible]

Query Match	Best Local Similarity	Matches 1117/	Conservative	0/	Mismatches	8/	Indels	86/	Gaps	1/
102	TTTTGAAGAGTGGGGTACGAACGCTTATCATATCATCTTGCTGACAGAAATGC	161	TTTTGAAGAGTGGGGTACGAACGCTTATCATATCATCTTGCTGACAGAAATGC	6581	TTTTGAAGAGTGGGGTACGAACGCTTATCATATCATCTTGCTGACAGAAATGC	6640				

QY	162	TCGACGAGGAAAGTCATCATAGGCAAGCATACACGAAACCTACGGCGAAAGACCTTCTCA	221
Db	6641	TGAACAGAGGAAAGTCATCATAGGCAAGCATACACGAAACCTACGGCGAAAGACCTTCTCA	6700
QY	222	GACTCTGCAAGGAGCTCTTAAGATTGG-----	253
Db	6701	GACTCTGCAAGGAGCTCTTAAGATTGG-----	6766
QY	254	-----AG	255
Db	6761	GACCTTTAAACATCTTGTTGAAGTCATTGTATTTCAATTTGTTTTTGGTCATGAGACG	6822
QY	256	AGAGCTATCTTGTGTGCACTTTGAACCCGGTAGCGGTATGCTTTATTTGGCTAATGA	315
Db	6821	AGAGCTATCTTGTGTGCACTTTGAACCCGGTAGCGGTATGCTTTATTTGGCTAATGA	6880
QY	316	GCTACAAAAGATGCACTTCAAGCAACCAAGTTCTTATGGAAGTTGCTTGCACAAGGCA	375
Db	6881	GCTACAAAAGATGCACTTCAAGCAACCAAGTTCTTATGGAAGTTGCTTGCACAAGGCA	6940
QY	376	TCAAAGCAGCTGCTTCAAGCTAGGCAAGCTTTCATGCTGCTGCACAAGAGTCTTTGA	435
Db	6941	TCAAAGCAGCTGCTTCAAGCTAGGCAAGCTTTCATGCTGCTGCACAAGAGTCTTTGA	7000
QY	436	GAGAGCGTTGCTCACCAACTACCGGTGACTTCAAGAAAGCTTTGGTCTTCTTGTTAC	495
Db	7001	GAGAGCGTTGCTCACCAACTACCGGTGACTTCAAGAAAGCTTTGGTCTTCTTGTTAC	7066
QY	496	TCATCAGGTTCGAAAGGAGATGCAATGGAACATGATGGCTTAAGCAAGAGCTTAAGCTG	555
Db	7061	TCATCAGGTTCGAAAGGAGATGCAATGGAACATGATGGCTTAAGCAAGAGCTTAAGCTG	7120
QY	556	GTCATGAGAAATCAAGGACAAAGCACTACATGATGAGATGTATTAGATCTTGTCC	615
Db	7121	GTCATGAGAAATCAAGGACAAAGCACTACATGATGAGATGTATTAGATCTTGTCC	7188
QY	616	ACAAGAGCAAAAGCTCAATCATCTTAAACGGTTACCAAGATGATCATAGGCGAG	675
Db	7181	ACAAGAGCAAAAGCTCAATCATCTTAAACGGTTACCAAGATGATCATAGGCGAG	7248
QY	676	GAAATTCGAAAGCTTGAAGAGAGATGATGACAAAGTTCCTTGCACTTTTGAG	735
Db	7241	GAAATTCGAAAGCTTGAAGAGAGATGATGACAAAGTTCCTTGCACTTTTGAG	7300
QY	736	TCAACCATTCAGTGCCTTGCAAGACAGAGCGTTTACTTTGTCAGATGCTTCGTGTCACA	795
Db	7301	TCAACCATTCAGTGCCTTGCAAGACAGAGCGTTTACTTTGTCAGATGCTTCGTGTCACA	7360
QY	796	ATCAACAAAACTGAACTGATGAAAGGACACTACTAGAAATTGTGACCAAGAGCTGAG	855
Db	7361	ATCAACAAAACTGAACTGATGAAAGGACACTACTAGAAATTGTGACCAAGAGCTGAG	7420
QY	856	ATTGACTTGAAGTCAATTGAGAGAGATACAGCGCAGAAACAGATTCCTTTGAGAAA	915
Db	7421	ATTGACTTGAAGTCAATTGAGAGAGATACAGCGCAGAAACAGATTCCTTTGAGAAA	7480
QY	916	GCTATTACCAAGACACTCGTAGAATTACGAGAAAGATGCTCGTGCACTTCGGGTAA	975
Db	7481	GCTATTACCAAGACACTCGTAGAATTACGAGAAAGATGCTCGTGCACTTCGGGTAA	7540
QY	976	GATGATGCTTAATCAATCAATCTTCCACAGAGAAACATTAAGCTGCTCTACAGCTTCTGT	1035
Db	7541	GATGATGCTTAATCAATCAATCTTCCACAGAGAAACATTAAGCTGCTCTACAGCTTCTGT	7600
QY	1036	ATCTCTTAATCTCCCTCTCTCTCTCTTTGATGAGTTCAATCGTTGATTTTGTGTTAC	1095
Db	7601	ATCTCTTAATCTCCCTCTCTCTCTCTTTGATGAGTTCAATCGTTGATTTTGTGTTAC	7660
QY	1096	AAAAACCTGTTTGTGTTGCTGTTGATTTGAGTTGCTTAATATGCAAAAGAGAGA	1155
Db	7661	AAAAACCTGTTTGTGTTGCTGTTGATTTGAGTTGCTTAATATGCAAAAGAGAGA	7720
QY	1156	CAGAGAGAACAGTGTGTCTCTTAAGTTATATATATGAGAGCATGTGCTTAATAA	1215

Db	7721	CAGAGAGAACGATGCTCTTAAAGTTATATATATGAAGCAATGGCTTAAACA	7780
QY	1216	AAAAAAAAA 1226	
Db	7781	CAGACTAACAA 7791	
RESULT 8			
CO806050			
LOCUS	CO806050	954 bp	DNA
DEFINITION	Sequence 2461 from Patent WO2004035798.		linear
ACCESSION	CO806050		
VERSION	CO806050.1	GI:47111659	
KEYWORDS			
SOURCE			
ORGANISM	Arabidopsis thaliana (chale cress)		
REFERENCE	1 Inze D., de Veylder, L. and Vlieghe, K.		
AUTHORS	Identification of novel e2f target genes and use thereof		
TITLE	Patent: WO 2004035798-A 2461 29-APR-2004;		
JOURNAL	CropDesign N.V. (BE)		
FEATURES	Location/Qualifiers		
source	1..954		
ORIGIN	/organism="Arabidopsis thaliana"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:3702"		
Query Match	77.6%; Score 954; DB 6; Length 954;		
Best Local Similarity	100.0%; Pred. No. 8e-231;		
Matches	954; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
QY	34	ATGGGCACTTAAAGTTCTTGATTCGTTCCTCTCTCTGATGATGCTGACCAATTG	93
Db	1	ATGGGCACTTAAAGTTCTTGATTCGTTCCTCTCTCTGATGATGCTGACCAATTG	60
QY	94	AGAACCGCTTTTGAAGATGGGGATACGAACGAGACTTGATCATATCAATCTTGGCTCAC	153
Db	61	AGAACCGCTTTTGAAGATGGGGATACGAACGAGACTTGATCATATCAATCTTGGCTCAC	120
QY	154	AGAAGTGCTGAACGAGGAAGTATATGACGCAACATATACCGAAGCTTACGGGGAAGAC	213
Db	121	AGAAGTGCTGAACGAGGAAGTATATGACGCAACATATACCGAAGCTTACGGGGAAGAC	180
QY	214	CTTCTCAAGACTCTTGAACAAGAGCTCTCTACATTTTCGAGAGAGCTATCTTGTGTGG	273
Db	181	CTTCTCAAGACTCTTGAACAAGAGCTCTCTACATTTTCGAGAGAGCTATCTTGTGTGG	240
QY	274	ACTCTTGAAACCGGTGAGCGTGATGCTTTATTTGGCTATGAAAGCTACAAAAGATGACT	333
Db	241	ACTCTTGAAACCGGTGAGCGTGATGCTTTATTTGGCTATGAAAGCTACAAAAGATGACT	300
QY	334	TCAAGCAACCAAGTTCTTATGGAAGTTGCTTGCACAAGACATCAAGCAGCTGCTTCAC	393
Db	301	TCAAGCAACCAAGTTCTTATGGAAGTTGCTTGCACAAGACATCAAGCAGCTGCTTCAC	360
QY	394	GCTAAGGCAAGTTACCAATGCTCGCTACAGAAGTCTCTTGAAGAAGAGCTTGCACAC	453
Db	361	GCTAAGGCAAGTTACCAATGCTCGCTACAGAAGTCTCTTGAAGAAGAGCTTGCACAC	420
QY	454	ACTACCGGTGACTTCAAGAAAGCTTTGGTCTTCTTGTATCTCATACAGTACGAAGA	513
Db	421	ACTACCGGTGACTTCAAGAAAGCTTTGGTCTTCTTGTATCTCATACAGTACGAAGA	480
QY	514	GATGAAGGCAACATGACATTTGGCTTAACAGAAGCTAAGCTGCTCCATGAGAAATTAAG	573
Db	481	GATGAAGGCAACATGACATTTGGCTTAACAGAAGCTAAGCTGCTCCATGAGAAATTAAG	540
QY	574	GACAAGCACTCATGATGAGATGTTATTAGATCTTGTCCACAAAGCAAGCTCAG	633

Db 541 GACCAAGCACTAACAAATGATGAGATGTTATTAGATCTTTTCACAAGAAAGCAAAAGCTCAG 600

Qy 634 ATCATGCTACTCTTTTAACCGTTACCAAGATGATCATGGGAGAAATTTCTCAAGAGCTCT 693

Db 601 ATCAATGCTACTCTTTTAACCGTTACCAAGATGATCATGGGAGAAATTTCTCAAGAGCTCT 660

Qy 694 GAGGAAGAGATGATGATGACAAAGTCTCTTGCACTTTTGAGTCAACCAATTCAGTCTTG 753

Db 661 GAGGAAGAGATGATGATGACAAAGTCTCTTGCACTTTTGAGTCAACCAATTCAGTCTTG 720

Qy 754 ACAAGACCAGAGCTTTTACTTTGTGATGTTCTTTCGTTACGCAATCAACAAACCTGGAAC 813

Db 721 ACAAGACCAGAGCTTTTACTTTGTGATGTTCTTTCGTTACGCAATCAACAAACCTGGAAC 780

Qy 814 GATGAAGGAGCACTCACTTGAAATTTGTGACCAACAAGCTGAGATTGACTTTGAAGGTCATT 873

Db 781 GATGAAGGAGCACTCACTTGAAATTTGTGACCAACAAGCTGAGATTGACTTTGAAGGTCATT 840

Qy 874 GGAGAGGAGTACCAAGGCGCAGAAACAGCAATTCCTTTGGAGAAAGCTATTACCAAGACACT 933

Db 841 GGAGAGGAGTACCAAGGCGCAGAAACAGCAATTCCTTTGGAGAAAGCTATTACCAAGACACT 900

Qy 934 CGTGAAGATTACGAGAAGATGCTCGTCGCACTTCTCGGTGAAGATGATGCTTAA 987

Db 901 CGTGAAGATTACGAGAAGATGCTCGTCGCACTTCTCGGTGAAGATGATGCTTAA 954

```

/clone="C00123 (e) "
/ecotype="Columbia"
/notes="This clone is now in pUNI 51 vector. Previously,
it was in pUNI-R3-D/V5-His-TOPO under the clone number
C00123. "
1. .954
/gene="At1g35720"
1. .954
/gene="At1g35720"
/codon_start=1
/evidence=experimental
/product="putative Ca2+-dependent membrane-binding protein
annexin"
/protein_id="AAC48798.1"
/db_xref="GI:12083278"
/translation="MATLKVSQSDVPAASDADAEQRTAFEGMGNEDLITSLHRSAB
QRKVIQRAHYHTYGEDLLKTLDKELSNDFERRALLWTLFEGEDDALANEAATKRWTS
NOYLMVEACRTSTGDLHARQAHYARFKSLSEEDVAHHTGDRKLLVSTSVRYGG
DEVNMTLAKOBAKLVHEKIKDKQHNDEDVIRLSTSKQAINTFNRYODDGEILK
SLEGGDDDDKFLALIRSTIOCLTRPELVPFDVARSANKTGDEGALTRLVTRAEID
LKVIQBEYQRRNSTPLEKAITKQTRGYEKMLVALGSDDA"

```

RESULT #					
AF332435		954 bp	mRNA	linear	PLN 07-MAY-200
LOCUS DEFINITION	Arabidopsis thaliana clone C00123 (e) putative Ca <sup>2+</sup> -dependent membrane-binding protein annexin (Atlg35720) mRNA, complete cds.				
ACCESSION	AF332435				
VERSION	AF332435.1 GI:12083277				
KEYWORDS	FLI CDNA.				
SOURCE ORGANISM	Arabidopsis thaliana (thale cress)				
REFERENCE AUTHORS	Arabidopsis thaliana Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids. 1 (bases 1 to 954) Yamada,K., Dale,J.M., Huan,V.W., Onodera,C.S., Quach,H.L., Toriumi,M., Wong,C., Wu,H.C., Yu,G., Yuan,S., Chen,H., Cheuk,R., Jones,T., Kim,C.J., Nguyen,M., Palm,C.J., Shinn,P., Southwick,A., Tripp,M.G., Wu,T., Davis,R.W., Ecker,J.R. and Theologis,A. Arabidopsis Open Reading Frame (ORF) Clones Unpublished 2 (bases 1 to 954) Yamada,K., Liu,S.X., Sakano,H., Pam,P.K., Banh,J., Ergu,P., Lee,J.M., Toriumi,M., Yu,G., Brooks,S., Chao,Q., Chen,H., Karlin-Neuman,G., Kim,C., Lam,B., Miranda,M., Nguyen,M., Palm,C.J., Shinn,P., Southwick,A., Davis,R.W., Ecker,J.R. and Theologis,A. Direct Submission Submitted (28-DEC-2000) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA 3 (bases 1 to 954) Yamada,K., Dale,J.M., Huan,V.W., Onodera,C.S., Quach,H.L., Toriumi,M., Wong,C., Wu,H.C., Yu,G., Yuan,S., Chen,H., Cheuk,R., Jones,T., Kim,C.J., Nguyen,M., Palm,C.J., Shinn,P., Southwick,A., Tripp,M.G., Wu,T., Davis,R.W., Ecker,J.R. and Theologis,A. Direct Submission Submitted (07-MAY-2003) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA Annotation based on July 2002 version of the Arabidopsis genome submitted to Genbank.				
TITLE JOURNAL COMMENT	Location/Qualifiers				
FEATURES	1..954				
source	/organism="Arabidopsis thaliana" /mol_type="mRNA" /db_xref="taxon:3702" /chromosome="1"				

OY	3*	ATGGGCACTCTTAAGTCTTTCGATCGTCGTCCTTCTGATGTAGTCGACAAATTG	93
Db	1	ATGGCGACTTTAAGGTTTCTGATTCGTCTCGCTCTTCTGTAGTATGCTGACCAATTG	60
OY	94	AGAACCCCTTTGAAGATGGGTACGACGAGGACTTGATCATATCAATCTTGGCTCAC	153
Db	61	AGAACCGCTTTTGAAGGATGGGGTATCGAACGAGGACTTGATCATATCTTGGCTCAC	120
OY	154	AGAACTGCTGAACGAGGAAAATCATAGCGACATATCCACGAACCTTACGGCGGAAGAC	213
Db	121	AGAACTGCTGAACGAGGAAAATCATAGCGACATATCCACGAACCTTACGGCGGAAGAC	180
OY	214	CTTCTCAAGACTCTTGAACAAGGACCTCTTAACGATTTGAGAGAGGCTATCTTGTGTGG	273
Db	181	CTTCTCAAGACTCTTGAACAAGGACCTCTTAACGATTTGAGAGAGCTATCTTGTGTGG	240
OY	274	ACTCTTGAACCCGGGTGAGCGTGAAGCTTTATTTGGCTATATGAAAGTCAAAAAAGATGACT	333
Db	241	ACTCTTGAACCCGGGTGAGCGTGAAGCTTTATTTGGCTATATGAAAGTCAAAAAAGATGACT	300
OY	334	TCAGCAACCAAGTTCTTATGGAAGTTGCTTGCACAAGGACATTAAGCGAGCTCTTCAC	393
Db	301	TCAGCAACCAAGTTCTTATGGAAGTTGCTTGCACAAGGACATTCMAAGCAGCTGCTTCAC	360
OY	394	GCTAGGCAAGTTTACATGCTCGCTACAAAGAAGCTCTTGAAGAGAGAGTGTGCTCACAC	453
Db	361	GCTAGGCAAGTTTACATGCTCGCTACAAAGAAGCTCTTGAAGAGAGAGTGTGCTCACAC	420
OY	454	ACTACCGGTGACTTCAGAAAAGCTTTTGTTCTTCTTGTATCCTCATACAGGTATCGAAGA	513
Db	421	ACTACCGGTGACTTCAGAAAAGCTTTTGTTCTTCTTGTATCCTCATACAGGTATCGAAGA	480
OY	514	GATGAAGTGAACATGACATTTGGCTTAAGCAAGAGAGCTTAAGCTGGTCCATGAGAAAATCAAG	573
Db	481	GATGAAGTGAACATGACATTTGGCTTAAGCAAGAGAGCTTAAGCTGGTCCATGAGAAAATCAAG	540
OY	574	GACAAGCACTATCATATGATGAGGATTAATTAAGATCTTGTTCACAAGAAGCAAAAGCTCAG	633
Db	541	GACAAGCACTATCATATGATGAGGATTAATTAAGATCTTGTTCACAAGAAGCAAAAGCTCAG	600
OY	634	ATCATATGCTACTTTTAAACCGTTTACCAAGATGATCATGGCGAGAAATTCCTCAAGAGCTTT	693
Db	601	ATCATATGCTACTTTTAAACCGTTTACCAAGATGATCATGGCGAGAAATTCCTCAAGAGCTTT	660
OY	694	GAGGAGGAGATGATGATGACAAAGTTCTTTCGACTTTTGAAGTCAACATTCAGTCTTG	753

QY	34	ATGGCGACTCTTAAAGGTTTCGATTCCTGCTCCTTGATGATGCTGAGCAATTG	93
Db	1	ATGGCGACTCTTAAAGGTTTCCTCTGCTCTCTCTCTGGAAGATGCTGAGCAATTG	60
QY	94	AGAACCGCTTTTGAAGGATGGGAGTACGAACGAGCATTTGATCATATCAATCTTGGCTCAC	153
Db	61	AAAAGCGGATTTGATGGATGGGAGTACCAACGAGGATTTGATCATATCAATCTTGGCTCAC	120
QY	154	AGAAAGTGTGAACGAGGAGAAAGTCAATCAGGCAAGATCAACAGAAACCTTACGGCGGAGAGC	213

SOURCE ORGANISM  
Gossypium hirsutum (upland cotton)  
Gossypium hirsutum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eustosids II; Malvales; Malvaceae; Malvoideae; Gossypium.  
1 (bases 1 to 1141)  
Shin, H. and Brown, R.M.Jr.  
Direct Submission  
Submitted (13-FEB-1997) Botany, The Univ. of Texas at Austin,  
Austin, TX 78713, USA



## FEATURES

CDS	Location/Qualifiers
source	1..1141 /organism="Goesyrium hirsutum" /mol_type="mrna" /strain="Texas Marker1" /db_xref="taxon:3635"
CDS	34..984

ORIGIN

Query Match	40.5%;	Score 497.8;	DB 8;	Length 1141;
Best Local Similarity	70.7%;	Pred. No. 3.8e-115;		
Matches 691; Conservative	0;	Mismatches 282;	Indels 4;	Gaps 2;

/transiaaiaa"mattllyppltyvpsvdscoqlaacsqmgllnbgllitlthgrm  
 onulirktvaytgebdlkadbeksndpberlvtlmaldpardmalilnbtikrwtss  
 novumeiactyrsanollhroaahaykssleedvahnthtgdfrllallpvs  
 eeuvnnlaktrealtheeksidraksoodvityrlarsaqolnaltlnhkneygd  
 tlakpodeftallbratvkclypkypekutalalnrrgdbealrtvvetraevdl  
 tlvadpqrnsuylprralvkcjohgeyexlllvtlavghevn

Qy	731	IGAGGTCAACCATTCAGTGGCTTGAACAGACAGACTTACTTTGTGCATGTTCTTCGTT	790
Db	728	TAAAGTCCACAGTAAAGTGGTCTTATCCGAAAGATATTTTGAAGGTTCTTCGCGC	787
Qy	791	CAGCAATCAACAAACTGGAATCGATGAAGAGCACTACTAGATTGTGACCAAGAG	850
Db	788	TAGCATCAATAGACAGAGAACGATGAAGAGCTCTTACTAGATTTGTTGTCATAGGG	847
Qy	851	CTGAGATTGACTTTGAAGTCAATTGGAGAGAGTACCTACGSCAGAGAACAGACTTCTTTGG	910
Db	848	CTGAGTTGATCTAAAGATCTAGCAGATGAGTACCAACGAAAGAACAGTGTCCACCTGA	907
Qy	911	AGAAAGCTATTACAAAGACACTCGTGGAGATTACGAAGAAGTCTCTCGCACTTTCTCG	970
Db	908	CTCGTGCCATTGTCAAGACACTCATGAGACTATGAAATAATGCTGTGCTGACTTTGCAG	967
Qy	971	GTGAAGATGATGCTTAA	987
Db	968	GACATGTGGAGAAATTGA	984

RESULT 12			
GHU73746	948 bp	mRNA	linear
LOCUS			PLN 24-JAN-2003
DEFINITION	<i>Gossypium hirsutum</i> amnecin (AmnG1) mRNA, partial cds.		

REFERENCE	1 (bases 1 to 948)
AUTHORS	Delmer,D.P. and Potikha,T.S.
TITLE	Structures and functions of annexins in plants
JOURNAL	Cell. Mol. Life Sci. 53 (6), 546-553 (1997)
MEDLINE	97374485
PubMed	9230934
REFERENCE	2 (bases 1 to 948)
AUTHORS	Potikha,T.S. and Delmer,D.P.
TITLE	Direct Submission
JOURNAL	Submitted (09-OCT-1996) Plant Sciences, Hebrew University, Givat Ram Campus, Jerusalem 91904, Israel
COMMENT	On Aug 20, 2001 this sequence version replaced gi.1843524.
FEATURES	Location/Qualifiers
source	1. .948

Oy	491	TTACCTCATCAGGATGAGGAAGGATGAAGTAAACATGCATTGGCTAAGCAAGAAGCTTA	550
Db	491	TGACTTCATCAGATATGAGGAGAGAGAGGTGAACATGCATCTGGCGAAAACAAGGCCA	550
Oy	551	AGCTGGCCATGAGAAANTCAAGNACAAGCATACATATGAGATGTTTATTAGAATCT	610
Db	551	AGTGCTTCATGAGAAAAATTTCAACAAGCTTACAGTAGTCAGATGTCATPAAGGTTT	610
Oy	611	TGTCACAAGAGCAAAGCTCAGATCAATGCTACTTTTAAACGGTTACCAAGATGCATC	670
Db	611	TGCGTAAAGAGGAAAGGACAAGATCAATGCAACTGTAATCATCTCAAAAAATGAATATG	670
Oy	671	GCGAGAAATTTCTCAAGAGCTTGAAGAGAGATGATGTCACAAGTTCTTGCACCTT	730
Db	671	GAATATGCATPAACAGAGACTT--GAAAGCTATCTTAAGATGAGTTCTTGGCACCTAC	727

gmsr...annexin

```

/function="evidence suggests that annexin may bind to
/influence activity of callose synthase"
/note="contains 4 repeats characteristic of other
annexins; shows evidence of conserved calcium-binding
domains; by comparison with other plant annexins as well
as with the size of native cotton fiber annexin, this
clone probably lacks 5'coding sequence for only the
N-terminal Met; calcium binding protein"
/codon_start=1
/product="annexin"
/protein_id="PAB67993.2"
/db_xref="GI:15214410"
/tran_slaction="ATLVPTVPSVEDCEOLRKAFGSGVTNEGIIIDILGRNAAE
RNLRKTYAVTEAGEDLLKALDELSNDPERFLVALADPEARDALLANAEATRWMSRS
QVIMEIACTRSANQLIHARQAYHARKYSLEBDVAHHHTGDPHKLILPVSYREGG
EVMVLIAKTAEAKLILHEKISNKAYSDDDVIRVIATATSKQINFTNLHYKNREYONDINKI

```

ORIGIN  
 LKADPKEFLALRSIVKCLVPEKYEKVLRLAINRGTDGALTRVCTRAEVDLK  
 VIADBYGRNRSVPLTRAIIVKDTGDIYKLLIVLAGEVA"

Query Match 39.7%; Score 488.6; DB 8; Length 948;  
 Best Local Similarity 70.3%; Pred. No. 8.3e-113;  
 Matches 669; Conservative 0; Mismatches 279; Indels 3; Gaps 1;

QY 37 GCCACTTTAAGTTTGTGATTCCTGCTCCTTCTGATGATGCTGACCAATTGAGA 96  
 DB 1 GCCACTTTAAGTTTGTGATTCCTGCTCCTTCTGATGATGCTGACCAATTGAGA 60  
 QY 97 ACCGCTTTGAAGGATGGGGTACGAGGAGGCTGATCAATCAATCTGGCTCAAGA 156  
 DB 61 AAGGCTTTTCAGATGGGGAATCAATGAGGCTTAACTAATGATTAATGGGTCACAGA 120  
 QY 157 AGTCTGAGAGAGAGAAAGTATCATAGGCAAGCATATACGAAACCTAGCGGAGACCTT 216  
 DB 121 AATGCGAGAGCAAGAACTTGATTCGAAAAACCTAGCGTGAACCTATGAGAGGATCTC 180  
 QY 217 CTCAGACTCTTGACAGAGAGCTCTCTAAGCATTTGAGAGAGCTATCTTGTGTGACT 276  
 DB 181 CTCAGGCACTAGACAGAACTCTCGAATGACTTTGAGAGGCTGTTCTGCTTGGGCT 240  
 QY 277 CTGGAACCGGTAGCGGTATGCTTTAATGCTTAATGAGCTTCAAAAAGATGACTTCA 336  
 DB 241 CTGATCTCTGAGAGCTGATGCTCTTGTGCTTAATAGCCACAAAGGTGACTTCA 300  
 QY 337 AGCAACCAAGTTCTTATGAGAGTTGCTTGACAAAGACATCAAGCAGCTGCTTCAAGT 396  
 DB 301 AGCATCAAGTCTCTTATGAGAAATAGCCTGCAAGAGCTGCTGCAAGCAAGCTTCAAGCA 360  
 QY 397 AGGCAAGCTTACCATGCTGCTTACAGAAAGCTTCTTGAAGAGAGCTTGTCTCACCACT 456  
 DB 361 AGGCAAGCTTATCATGCTGCTTATGAAGATCGCTTGAAGAGAGCTTGTCTCATACAGC 420  
 QY 457 ACCGCTGACTTCAAGAGCTTTTGTGTTCTCTGTTACCTATACAGTACGAAAGAGAT 516  
 DB 421 ACTGGGAGCTTCCATTAAGCTCTCTCTTACCTTGTAGTTCATACAGATATGAGGGAG 480  
 QY 517 GAAGTGAACATGATCATTTGCTTAAAGCAAGAGCTTGTGCTCAAGAAATCAAGAGC 576  
 DB 481 GAGGTGAACATGATCTGCGAAAGAGAGAGAGAGCTTGTGCTCAAGAAATTTTCAAC 540  
 QY 577 AAGCACTACATGATGAGATGTTATTAAGATCTTGTCAAGAGAGCAAGCTCAGATC 636  
 DB 541 AAGCTTACATGATGAGATGATCAATAGGATTTGGCTCAAGAGAGCAAGCTCAGATC 600  
 QY 637 AATGCTACTTTTAAACCTTACCAAGATGATCAGGAGAGAAATCTCAAGAGCTTGGAG 636  
 DB 601 AATGCAACTTGATCTCAAAAGATGATGAGAAATGAGAAATGAGAAAGCTTGAAG 660  
 QY 697 GAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 756  
 DB 661 GCTGACCTTAAGATG---AGTTCTTGTGCTACTAGTAAAGTGTGCTGCT 717  
 QY 757 AGACCAAGGCTTACTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 816  
 DB 718 TATCCGGAAGATATTTGAGAGGTTCTTCCGCTAGCATATAGACGAGAGCGAT 777  
 QY 817 GAAGAGCACTACTAGATTTGATGACCAAGAGCTGATGATGATGATGATGATGATGATGAT 876  
 DB 778 GAAGAGCTTCTTCAAGATGTTTGTGATGAGGCTGAGGCTGATCTAAAGGCTCAAGCA 837  
 QY 877 GAGAGATACGAGAGAGAGAGAGATCTTGTGAGAAAGCTATTAACAAAGACATCGT 926  
 DB 838 GATGAGATACGAGAGAGAGAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 897  
 QY 937 GAGATTAAGAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 987  
 DB 898 GGAGACTATGAAAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 948

## RESULT 13

AF006197

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCES

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCES

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCES

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCES

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCES

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCES

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCES

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCES

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCES

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCES

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCES

AUTHORS

TITLE

JOURNAL

AF006197 1112 bp mRNA linear PLN 22-AUG-2001  
 Lavatera thuringiaca annexin (AnxLc1) mRNA, complete cds.

AF006197.1 GI:2459925

Lavatera thuringiaca

Lavatera thuringiaca

Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

Rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Lavatera.

1 (bases 1 to 1112)

Vazquez-Tello, A. and Uozumi, T.

Cloning and characterization of a Lavatera thuringiaca cDNA

encoding an annexin whose expression is stimulated by low

temperature

Plant Physiol. (1997) In press

2 (bases 1 to 1112)

Breton, G., Vazquez-Tello, A., Danyluk, J. and Sarhan, F.

Two novel intrinsic annexins accumulate in wheat membranes in

response to low temperature

Plant Cell Physiol. 41 (2), 177-184 (2000)

20255875

10795312

3 (bases 1 to 1112)

Vazquez-Tello, A. and Uozumi, T.

Submitted (31-MAY-1997) Biological Sciences, Universite du Quebec,

Montreal, C.P. 8888, Succ. Centre-Ville, Montreal, Quebec H3C 3J9,

Canada

## FEATURES

source

1..1112

/organism="Lavatera thuringiaca"

/mol\_type="mRNA"

/db\_xref="taxon:61660"

/feature\_type="leaves without petioles"

/note="From England"

1..1112

/gene="AnxLc1"

2..952

/gene="AnxLc1"

/function="binds acidic phospholipids in a

calcium-dependent manner"

/note="calcium-dependent, phospholipid binding protein;

contains four characteristic repeats found in both animal

and plant annexins"

/codon\_start=1

/product="annexin"

/protein\_id="AAB71830.1"

/db\_xref="GI:2459926"

/translation="MATLVTPSLPSVSDRCQLRKAFGMTNEDLIIINLGHNRAD

ENSTIRKATETHGDLKALDKELSNDEPRVILMTDPPRDLINAEATRWTS

NOVIMEIRSSSDILRQARQAYHRYKSLSEVDVAHHTGPRKLLIPVSYRIG

DRVNNLTAKTEKLHEKISNRAYSDDVDVTLATRSQSINERLNHTNRYATDINK

DKADPKDFLALRSIVKCLVPEKYEKVLRLAINRGTDGALTRVSTRAEVDL

KIIADBYGRNRSVPLTRAIIVKDTGDIYKLLIVLAGEVA"

## ORIGIN

Query Match 38.9%; Score 478.6; DB 8; Length 1112;

Best Local Similarity 69.5%; Pred. No. 2.8e-110;

Matches 665; Conservative 0; Mismatches 289; Indels 3; Gaps 1;

QY 33 AATGGCACTTTAAGTTTGTGATTCCTGCTCCTTCTGATGATGCTGACCAATT 92  
 DB 1 AATGGCTACTTTACAGTTCCCTCCACACTTCCGTCAGTGTGAAAGTGAACACT 60  
 QY 93 GAGAACCGCTTTTGAAGATGGGGTACGAGAGAGCTTATCATATCATCTTGGCTCA 152  
 DB 61 CAGGAAGCTTCTCAGATGGGGAATGATGAGAGCTTAAATATATATTTGGGTCA 120  
 QY 153 CAGAGTGTGAACAGAGAAAGTATCATGAGAAACATATCAAGAACTTAGGGGAGA 212

Db 121 CCGAATGCGGACGAAAGAACTCGATTGCGAAAGCTTACTGAAAGCCATGAGAGAGA 180  
QY 213 CCTTTCAGAGCTCTTGACAAAGAGCTCTTACGATTTGAGAGAGCTATCTGTGTG 272  
Db 181 TCTCTTCAAGGCACTGACAGAACTCTTCAATGACTTTGAGAGGCTGTCTGTG 240  
QY 273 GACTCTTGAACCCGCTGAGCGTGTCTTTATTTGGCTAAAGAACTCAAAAAGATGAC 332  
Db 241 GACTCTTGAATCTCTGGAAGTATGATGACTTTTGGCAATGAGCCACAAAAGGTGAC 300  
QY 333 TTCAACCAACCAAGTTCTTATGGAAGTGTCTGACAAAGACTCAACGAGTGTCTTA 392  
Db 301 TTCACCAATCAGATTAATTAAGAAATAGCTGACAAAGTCTTCTGACCACTGCTTG 360  
QY 393 CGGTAGGACGCTTACATGCTGCTACAGAAAGTCTCTTGAAGAGAGCTGTGCTACCA 452  
Db 361 CCGGAGGACGCTTACATGCTGTCTTATGAAATGCTTGAAGAGATGTTGCCATCA 420  
QY 453 CACTACCGGTGACTTACAGAAAGCTTTGTTCTCTGTTACCTCATACAGGTACGAG 512  
Db 421 CACAACTGGGCACTTCGTAAGCTTCTCTACTCTTGTGAGTTATACAGATACGAGG 480  
QY 513 AGATGAAGTGAATGACATGCTGCTTACAGAAAGCTAGCTGTCTCATGAGAAATCA 572  
Db 481 AGATGAAGTGAATGACATGCTGCTGCAAAAACAGAGGCAAGTTACTCATGAGAAATCTC 540  
QY 573 GGAACAAGCACTACATGATGAGATGTTATTAAGATCTTCCACAAGAGCAAAAGCTCA 632  
Db 541 AAACAAGCTTACAGATGAGATGATGCTCATGAGGTTTGGCTACAGAAAGCAAGTCA 600  
QY 633 GATCAATGCTACTTTTACCGTTACCAAGATGATCATGCGAGAAATTTCTCAAGTCT 692  
Db 601 GATCAACGAACGCTTATATCACTACAAATGAAATGACCAACTGATATTAACAAAGACCT 660  
QY 693 TGAAGAGAGATGATGATGACAAAGTCTCTTGAATTTGAGTCAACCATTCAGCTT 752  
Db 661 GAAGGCTGACCTTAAGATG--AGTTCCTGACATGCTAAGGTCCACAGTGAAGTCTT 717  
QY 753 GACAAAGACCAAGCTTACTTGTGATGTTCTGCTGACCAATCAACAAATCTGAAC 812  
Db 718 GGTCTACCTGAAAGATATTTGAGAAAGTCTTCTGCTAGCAATCAATTAACGAGAAC 777  
QY 813 TGAAGAGAGCACTGACTGAATTTGTGACCAACAGAGCTGAGATTGACTTGAAGTCT 872  
Db 778 GGTGAGAGAGCTCTTACGAGGTTGTTTCCACAGGCTGAGGTTGATCTAAAGTCTAT 837  
QY 873 TGAAGAGAGTACCAACGACAGAAACGATCTCTTGAAGAAAGCTATTAACAAAGAC 932  
Db 838 AGCAGATGAGTACCAACGAGAAACAGTGTCCACAGCTGCTGCTATTTGTCAAGGAC 897  
QY 933 TCGTGAAGATTACGAGAGATGCTGCTGCACTTCTCGGTGAAGATGATGCTTAATC 989  
Db 898 TATGAGAGACTAGAAATAATGCTGCTGATCTTCTGAGAGAGGTGAGAGGCTTGAAC 954

RESULT 14  
BD236006 2588 bp DNA linear PAT 17-JUL-2003  
LOCUS BD236006  
DEFINITION Materials and method for modification of plant cell wall  
polyaccharides.  
ACCESSION BD236006  
VERSION BD236006.1 GI:33045776  
KEYWORDS JP 2002527056-A/28.  
SOURCE  
ORGANISM  
Eucalyptus grandis  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; Myrtales; Myrtaceae; Eucalyptus.  
1 (bases 1 to 2588)  
Blokberg L.N.  
REFERENCE  
AUTHORS Materials and method for modification of plant cell wall  
TITLE Patent: JP 2002527056-A 28 27-AUG-2002;  
JOURNAL GENESIS RESEARCH AND DEVELOPMENT CORP LTD, FLETCHER CHALLENGE

COMMENT  
FORESTS LTD  
OS Eucalyptus grandis (flooded gum)  
PN JP 2002527056-A/28  
PD 27-AUG-2002  
PF 06-OCT-1999 JP 2000575985  
PR 13-OCT-1998 US 09/170862,11-AUG-1999 US 60/148426 PI  
LEONARD NATHAN BLOKBERG  
PC C12N15/00,A01H5/00,C12N5/10,C12N9/00,C12N9/12,C12N15/00,C12N5/  
PC 00  
CC Materials and method for modification of plant cell wall CC

FEATURES  
source location/Qualifiers  
FT 1..2588  
FT source location/Qualifiers  
FT 1..2588  
FT /organism='Eucalyptus grandis'  
/mol\_type='genomic DNA'  
/db\_xref='taxon:71139'

ORIGIN  
Query Match 38.4%; Score 471.8; DB 6; Length 2588;  
Best Local Similarity 68.6%; Pred. No. 1.4e-108;  
Matches 666; Conservative 0; Mismatches 302; Indels 3; Gaps 1;

QY 30 AAAATGCGCACTTAAAGTCTTCTGATCTGTTCTCTCTCTCTGATGATGCTGACCA 89  
Db 1386 AACATGTCGACTCTCAACGCTCCGACAGCACTGCCCCCTGATGCCATGACTGAGACA 1445  
QY 90 ATTGAACCCGCTTTGAAGATGGGGTACGACAGAGACTTGTATCATATCTTGGC 149  
Db 1446 GCTCCGACAGCTTTCGAGATGGGGAAACAATGAGAGGTGATCATATCATATTTGG 1505  
QY 150 TCACAGAGTGTCAACAGAGAAATCATAGGCAACATACACAGAACTTACAGGGA 209  
Db 1506 TCATGGAATGGCGGACAGAGAGACTGATTCGAAACCTATGCGGAGCTTACAGGGA 1565  
QY 210 AGACTTCTCAAGACTTGTGACAGAGACTCTTACAGATTTGAGAGGCTATCTTGT 269  
Db 1566 GACTTCAAGCAACGATTTGACAGAGAACTTACCAATGATTTCCAGAGGCTGTGCTCT 1625  
QY 270 GTGACTTTGAACCCGCTGAGCTGATGCTTATTTGCTATGAGAGCTTCAAAAAGTG 329  
Db 1626 TTGCTACCTTGTTCGCTGAAACGATGCTTACTTGGCAATGAGAGAGAAAGATG 1685  
QY 330 GACTTCAAGCAACCAAGTCTTATGAGAGTGTGCAAGAGCATCAAGCAGCTGCT 389  
Db 1686 GACTTCAAGCAACCAAGTCTTATGAGAAATGCTGCAAGAGTCTCCGACAGCTTGTCT 1745  
QY 390 TCACGCTGAGCAAGCTTACATGCTGCTACAGAAAGTCTCTTGAAGAGAGCTTGTCTCA 449  
Db 1746 TATGCAAGAGACAGATATCATGCCCCGATACAGAAAGTCTGAGAAAGAGAGCTGCTCA 1805  
QY 450 CCACACTACCGGTGACTTCAAGAAAGCTTTGGTCTTCTTGTATCTATACAGGTACGA 509  
Db 1806 CCACAACTGAGAGATTTTCTGTAAGTGTCTGTACTCTTGTGAGCTCTCTCCGTTTGA 1865  
QY 510 AGGATATGAGGAACATGATGCTGCTTACCAAGAGCTTAAAGCTGCTGCTAGAAAT 569  
Db 1866 TGAATATGAGGATATGATCTTTGGCAAAAGCAAGAGCTTAAAGTCTCCAGAGAAAT 1925  
QY 570 CAAGAGCAAGCACTCAATGATGAGATGTTATTAAGATCTTGTCCACAAAGCAAAAGC 629  
Db 1926 CTCAGAGAGGCTTATGCGCATGAGAGATCTCATAGATTTTGGCTATACAGAGCAAAAGC 1985  
QY 630 TCAGATCAATGCTATTTTAAACGCTTACCAAGATGATGAGCGAGAAATTTCTCAAGAG 689  
Db 1986 ACAGGTCAATGCTACGCTGAATCACTACAAAATGATTTGGAATGATATCAACAGGA 2045  
QY 690 TCTTGAAGAGAGATGATGATGACAAAGTCTTGTGAGTCACTTTTGAAGTCAACATTCAGTG 749  
Db 2046 TTT---GAAACTGATCCAAAGAGCGGTCTTACTATCTACTGAGAGCTTACGATTAAGTG 2102



**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using bw model

Run on: August 22, 2005, 03:05:29 ; Search time 805 Seconds  
(without alignments)  
9045.071 Million cell updates/sec

Title: US-10-690-564-1

Perfect score: 1230

Sequence: 1 ccacgcgctccgaacactaa.....aaaaaaaaaaaaaaaa 1230

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4390206 segs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_16Dec04:\*  
1: geneseqn1808:\*  
2: geneseqn1908:\*  
3: geneseqn2008:\*  
4: geneseqn2001as:\*  
5: geneseqn2001bs:\*  
6: geneseqn2002as:\*  
7: geneseqn2002bs:\*  
8: geneseqn2003as:\*  
9: geneseqn2003bs:\*  
10: geneseqn2003cs:\*  
11: geneseqn2003ds:\*  
12: geneseqn2004as:\*  
13: geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1162.4	94.5	1192	3	AAC34418	Aac34418 Arabidops
2	1144.8	93.1	1212	3	AAC34549	Aac34549 Arabidops
3	1141.4	92.8	1182	3	AAC33439	Aac33439 Arabidops
4	954	77.6	954	12	AD63871	Adj63871 Plant lip
5	954	77.6	954	12	ADN74566	Adn74566 Thale cre
6	635	51.6	635	6	AB06078	Abg66078 Arabidops
7	497.8	40.5	1184	13	ADB63461	Adt63461 Cotton CD
8	473.4	38.5	1058	10	ADB94854	Adb94854 Programme
9	471.8	38.4	2588	3	AAAF7100	Aaa67100 Eucalyptu
c 10	470	38.2	487	10	ADB81194	Aad81194 Arabidops
11	468.6	38.1	1293	4	AAAF4756	Aaf44756 Annexin-1
12	468.6	38.1	1293	4	AAAF4756	Aaf44756 Annexin-1
13	468.6	38.1	1293	4	AAAF4756	Aaf44756 Annexin-1
14	457.2	37.2	1160	3	AAC90038	Aac90038 Arabidops
15	455.2	37.0	1156	3	AAC9728	Aac9728 Arabidops
16	453.2	36.8	954	12	ADN74790	Adn74790 Thale cre
17	449.6	36.6	1157	10	ADP37996	Adp37996 Synchro
18	426	34.6	951	3	AAC41736	Aac41736 Arabidops
19	424	34.5	1025	3	AAC49738	Aac49738 Arabidops
20	413.6	33.6	1027	3	AAC37764	Aac37764 Arabidops

ALIGNMENTS

RESULT 1	21	364.2	29.6	789	3	AAAF7131	AAAF7131 Pinus rad
AAC34418	22	364.2	29.6	789	3	AAAF7098	Aaf67098 Eucalyptu
ID AAC34418 standard; DNA; 1192 BP.	23	364.2	29.6	789	4	AAAF4759	Aaf44759 Annexin-1
AC AAC34418;	24	364.2	29.6	789	10	ADB94946	Adb94946 Programme
XX	25	347.8	28.3	420	3	AAC37667	Aac37667 Arabidops
XX	26	327.2	26.6	602	13	ACN57334	Acn57334 Cotton gy
XX	27	327	26.6	591	13	ACN60265	Acn60265 Cotton gy
XX	28	326	26.5	589	13	ACN58684	Acn58684 Cotton gy
XX	29	322.2	26.2	600	13	ACN60840	Acn60840 Cotton gy
XX	30	315.8	25.7	586	13	ACN60745	Acn60745 Cotton gy
XX	31	294.2	23.9	954	12	ADQ07773	Adq07773 Polynucle
XX	32	282.6	23.0	570	13	ACN58229	Acn58229 Cotton gy
XX	33	271	22.0	608	13	ACN53624	Acn53624 Cotton an
XX	34	270.6	22.0	2132	3	AAAF7099	Aaf67099 Pinus rad
XX	35	270.4	22.0	454	3	ACN36827	Aac36827 Arabidops
XX	36	266.2	21.6	618	13	ADB63464	Adt63464 Cotton CD
XX	37	230.8	18.8	504	13	ACN59774	Acn59774 Cotton gy
XX	38	213.4	17.3	437	12	ADP95395	Adp95395 Cotton ex
XX	39	210	17.1	704	4	AAAF4760	Aaf44760 Annexin-1
XX	40	208.4	16.9	704	3	AAAF7132	Aaf67132 Pinus rad
XX	41	208.4	16.9	704	10	ADB94947	Adb94947 Programme
XX	42	206.8	16.8	704	3	AAAF7096	Aaf67096 Pinus rad
XX	43	198	16.1	396	12	ADP93369	Adp93369 Cotton ex
XX	44	186.2	15.1	484	3	AAAF7129	Aaf67129 Pinus rad
XX	45	186.2	15.1	484	4	AAAF4757	Aaf44757 Annexin-1
OS Arabidopsis thaliana.							
XX EP1033405-A2.							
XX PD 06-SEP-2000.							
XX PF 25-FEB-2000; 2000BP-00301439.							
XX PR 25-FEB-1999; 99US-0121825P.							
XX PR 05-MAR-1999; 99US-0123180P.							
XX PR 09-MAR-1999; 99US-0123548P.							
XX PR 23-MAR-1999; 99US-0125788P.							
XX PR 25-MAR-1999; 99US-0126264P.							
XX PR 29-MAR-1999; 99US-0126785P.							
XX PR 01-APR-1999; 99US-0127462P.							
XX PR 06-APR-1999; 99US-0128234P.							
XX PR 08-APR-1999; 99US-0128714P.							
XX PR 16-APR-1999; 99US-0128845P.							
XX PR 19-APR-1999; 99US-0130077P.							
XX PR 21-APR-1999; 99US-0130449P.							
XX PR 23-APR-1999; 99US-0130510P.							
XX PR 23-APR-1999; 99US-0130891P.							
XX PR 28-APR-1999; 99US-0131449P.							
XX PR 30-APR-1999; 99US-0132048P.							
XX PR 30-APR-1999; 99US-0132407P.							
XX PR 04-MAY-1999; 99US-0132484P.							
XX PR 05-MAY-1999; 99US-0132485P.							
XX PR 06-MAY-1999; 99US-0132486P.							
XX PR 06-MAY-1999; 99US-0132487P.							

PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0133256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 26-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 30-AUG-1999; 99US-0151080P.  
PR 31-AUG-1999; 99US-0151303P.  
PR 01-SEP-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156566P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.



```
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161933P.
PR 29-OCT-1999; 99US-0162142P.

Query Match      94.5%; Score 1162.4; DB 3; Length 1192;
Best Local Similarity 99.1%; Pred. No. 1,2e-286;
Matches 1169; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 10 CGAAGACTTAAAGTGAAGAAAAATGGCGACTCTTAAGGTTCTGATTCCTGCTGCT 69
DB 12 CAAAACACTAAAGTGAAGAAAAATGGCGACTCTTAAGGTTCTGATTCCTGCTGCT 71
QY 70 CCTTCTGATGATGCTGAGCAATTGAAACCGCTTTTGAAGATGGGGTACGAACGAGAC 129
DB 72 CCTTCTGATGATGCTGAGCAATTGAAACCGCTTTTGAAGATGGGGTACGAACGAGAC 131
QY 130 TTGATCATATCATCTTTGGCTCAGAGTGTCTGAACAGAGGAAGTCAATCAGGCAAGCA 189
DB 132 TTGATCATATCATCTTTGGCTCAGAGTGTCTGAACAGAGGAAGTCAATCAGGCAAGCA 191
QY 190 TACCAAGAAACCTAAGCGGCAAGACCTTCTCAAGACTCTTGACAAGAGGCTCTTAACGAT 249
DB 192 TACCAAGAAACCTAAGCGGCAAGACCTTCTCAAGACTCTTGACAAGAGGCTCTTAAACGAC 251
QY 250 TTGAGAGAGCTATCTTTGTTGTGACTCTTGAACCCGGTAGCGGTATGCTTATTTGCT 309
DB 252 TTGAGAGAGCTATCTTTGTTGTGACTCTTGAACCCGGTAGCGGTATGCTTATTTGCT 311
QY 310 AATGAAGCTTCAAAAAGATGAGATTTCAAGCAACCAAGTTCTTATGGAAGTTGCTTCAACA 369
DB 312 AATGAAGCTTCAAAAAGATGAGATTTCAAGCAACCAAGTTCTTATGGAAGTTGCTTCAACA 371
QY 370 AGGACATCAACGAGCTGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCT 429
DB 372 AGGACATCAACGAGCTGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCT 431
QY 430 CTGGAAGAGAGCGTTGCTCAACCACTACCGGTGACTTGAAGCTTTTGTGTTCTCTT 489
DB 432 CTGGAAGAGAGCGTTGCTCAACCACTACCGGTGACTTGAAGCTTTTGTGTTCTCTT 491
QY 490 GTTACTCTCATACGATGATGAGAGATGAAATGATGATGATGATGATGATGATGATGATGAT 549
DB 492 GTTACTCTCATACGATGATGAGAGATGAAATGATGATGATGATGATGATGATGATGATGAT 551
QY 550 AAGCTGCTCATGAGAAAAATCAAGACCAAGCACTAATGATGAGAGATGATGATGATGATGAT 609
DB 552 AAGCTGCTCATGAGAAAAATCAAGACCAAGCACTAATGATGAGAGATGATGATGATGATGAT 611
QY 610 TTGTCCACAAGAGCAAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 669
DB 612 TTGTCCACAAGAGCAAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 671
QY 670 GGGGAGGAATTTCTCAAGAGCTTTGAGAGAGAGATGATGATGATGATGATGATGATGAT 729
DB 672 GGGGAGGAATTTCTCAAGAGCTTTGAGAGAGAGATGATGATGATGATGATGATGATGAT 731
QY 730 TTGAGGTCAACCTTCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 789
DB 732 TTGAGGTCAACCTTCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 791
QY 790 TCAGCAATCAACAAAACCTGGAACCTGATGAGAGCACTCAATGATGATGATGATGATGAT 849
DB 792 TCAGCAATCAACAAAACCTGGAACCTGATGAGAGCACTCAATGATGATGATGATGATGAT 851
```

```
QY 850 GCTGAGATTGACTTGAAGTCAATTGAGAGAGTACACGCGAGGAACGATTCCTTTG 909
DB 852 GCTGAGATTGACTTGAAGTCAATTGAGAGAGTACACGCGAGGAACGATTCCTTTG 911
QY 910 GAGAAAGCTATTACCAAGAGACACTGGAGATTACGAGAAAGATGCTGTCGCACTTCTC 969
DB 912 GAGAAAGCTATTACCAAGAGACACTGGAGATTACGAGAAAGATGCTGTCGCACTTCTC 971
QY 970 GGTGAAGATGATGCTTAAATCAATCAATCTCCACAGAGAAACATTAAGTCTCTACAGCT 1029
DB 972 GGTGAAGATGATGCTTAAATCAATCAATCTCCACAGAGAAACATTAAGTCTCTACAGCT 1031
QY 1030 TCTGTTATCTTAAATCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1089
DB 1032 TCTGTTATCTTAAATCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1091
QY 1090 TTCTACAAAACCTTGTGTTGTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1149
DB 1092 TTCTACAAAACCTTGTGTTGTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1151
QY 1150 GAGAGACAGAGAAACCAAGTGTGTTCTTAAATTATATA 1189
DB 1152 GAGAGACAGAGAAACCAAGTGTGTTCTTAAATTATATA 1191

RESULT 2
AAC34549
ID AAC34549 standard; DNA; 1212 BP.
XX
AC AAC34549;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 7037.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
XX
PR 05-MAR-1999; 99US-0123180P.
XX
PR 09-MAR-1999; 99US-0123548P.
XX
PR 23-MAR-1999; 99US-0125788P.
XX
PR 25-MAR-1999; 99US-0126264P.
XX
PR 29-MAR-1999; 99US-0126785P.
XX
PR 01-APR-1999; 99US-0127462P.
XX
PR 06-APR-1999; 99US-0128234P.
XX
PR 08-APR-1999; 99US-0128714P.
XX
PR 16-APR-1999; 99US-0129845P.
XX
PR 19-APR-1999; 99US-0130077P.
XX
PR 21-APR-1999; 99US-0130449P.
XX
PR 23-APR-1999; 99US-0130510P.
XX
PR 23-APR-1999; 99US-0130891P.
XX
PR 28-APR-1999; 99US-0131449P.
XX
PR 30-APR-1999; 99US-0132048P.
XX
PR 30-APR-1999; 99US-0132407P.
XX
PR 04-MAY-1999; 99US-0132484P.
XX
PR 05-MAY-1999; 99US-0132485P.
XX
PR 06-MAY-1999; 99US-0132486P.
XX
PR 06-MAY-1999; 99US-0132487P.
XX
PR 07-MAY-1999; 99US-0132863P.
XX
PR 11-MAY-1999; 99US-0134256P.
XX
PR 14-MAY-1999; 99US-0134218P.
XX
PR 14-MAY-1999; 99US-0134219P.
```

PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139545P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142820P.  
PR 12-JUL-1999; 99US-0142877P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.

PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 03-AUG-1999; 99US-0146389P.  
PR 04-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 23-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0158293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.



PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-014091P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.

PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147415P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.



```
QY 34 ATGGGACCTCTTAAGGTTTCTGATTCGTTCCGTCCTTGTGATGATGAGCAATTG 93
DB 1 ATGGGACCTCTTAAGGTTTCTGATTCGTTCCGTCCTTGTGATGATGAGCAATTG 60
QY 94 AGAACCGCTTTTGAAGATGGGGTACGAAGAGACTTGTATCATATCAATCTTGGCTCAC 153
DB 61 AGAACCGCTTTTGAAGATGGGGTACGAAGAGACTTGTATCATATCAATCTTGGCTCAC 120
QY 154 AGAAGTCTGAAACAGAGAAAGTCATCAGGCAAGCATACCAAGAACTTACGGCCGAAGAC 213
DB 121 AGAAGTCTGAAACAGAGAAAGTCATCAGGCAAGCATACCAAGAACTTACGGCCGAAGAC 180
QY 214 CTTCTCAAGACTTTTGAACAGAGACTCTCTCAAGATTTTGAAGAGACTTCTTGTGTGG 273
DB 181 CTTCTCAAGACTTTTGAACAGAGACTCTCTCAAGATTTTGAAGAGACTTCTTGTGTGG 240
QY 274 ACTCTTGAACCCCGTAGCGGTGATGCTTTATTTGGCTTATGAAAGCTACAAAAGATGAGCT 333
DB 241 ACTCTTGAACCCCGTAGCGGTGATGCTTTATTTGGCTTATGAAAGCTACAAAAGATGAGCT 300
QY 334 TCAGCAACCAAGTTCTTATGAAAGTTGCTTGCAAGACATCAACGAGCTGCTTCAAC 393
DB 301 TCAGCAACCAAGTTCTTATGAAAGTTGCTTGCAAGACATCAACGAGCTGCTTCAAC 360
QY 394 GCTAAGCAAGCTTACAGATGCTGCTCAAGAAAGCTCTTGAAGAGAGCGTTGCTCACAC 453
DB 361 GCTAAGCAAGCTTACAGATGCTGCTCAAGAAAGCTCTTGAAGAGAGCGTTGCTCACAC 420
QY 454 ACTACCGGTGACTTCAGAAAGCTTTGGTTCTCTTGTATACCTATACAGATACGAAGA 513
DB 421 ACTACCGGTGACTTCAGAAAGCTTTGGTTCTCTTGTATACCTATACAGATACGAAGA 480
QY 514 GATGAAGTGAACATGACATTTGGCTAAGCAAGAAAGCTTGTCCATGAGAAATCAG 573
DB 481 GATGAAGTGAACATGACATTTGGCTAAGCAAGAAAGCTTGTCCATGAGAAATCAG 540
QY 574 GACAGCACTACATGATGATGAGATGTTATGAGATCTTGGCCCAAGAAAGCAAGAGCTGAC 633
DB 541 GACAGCACTACATGATGATGAGATGTTATGAGATCTTGGCCCAAGAAAGCAAGAGCTGAC 600
QY 634 ATCAATGCTACTTTTAAACGTTTACCAAGATGATCATGGCGAGAAATTTCTCAAGACTTT 693
DB 601 ATCAATGCTACTTTTAAACGTTTACCAAGATGATCATGGCGAGAAATTTCTCAAGACTTT 660
QY 694 GAGGAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 753
DB 661 GAGGAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 720
QY 754 ACAAGACGAGAGCTTACTTTCGATGATGATGATGATGATGATGATGATGATGATGATG 813
DB 721 ACAAGACGAGAGCTTACTTTCGATGATGATGATGATGATGATGATGATGATGATGATG 780
QY 814 GATGAAGGACCTCACTAGAAATTTGACCAAGAGCTGAGATTTGATTTGAAGATCAT 873
DB 781 GATGAAGGACCTCACTAGAAATTTGACCAAGAGCTGAGATTTGATTTGAAGATCAT 840
QY 874 GAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 933
DB 841 GAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 900
QY 934 CGTGAAGATTAAGAGAAAGATGCTGCTGCACTTCTGCGGTGAAGATGATGATGAT 987
DB 901 CGTGAAGATTAAGAGAAAGATGCTGCTGCACTTCTGCGGTGAAGATGATGATGAT 954
```

```
RESULT 5
ID ADN74566 standard; cDNA; 954 BP.
AC ADN74566;
XX
DT 15-JUL-2004 (first entry)
```

```
XX DE Thale cress cDNA repressed in E2Fa/Dpa expressing plants Segid 2461.
XX KM gene; ss; plant; transgenic; E2Fa/Dpa transcription factor;
XX KW growth regulator; animal feed product; thale cress;
XX OS cell wall biosynthesis; nitrogen metabolism; carbon metabolism.
XX Arabidopsis thaliana.
XX PN WO2004035798-A2.
XX PD 29-APR-2004.
XX PF 20-OCT-2003; 2003WO-EP011658.
XX PR 18-OCT-2002; 2002EP-00079408.
XX (CROP-) CROPDISEIGN NV.
XX Inze D, De Veylder L, Vlieghe K;
XX WPI; 2004-348466/32.
XX DR P-PSDB; ADN74567.
XX PT Altering plant characteristics, useful for producing plants for enzyme or
XX PT pharmaceutical production comprises modifying in a plant, expression of
XX PT one or more nucleic acids and/or modifying level or activity of one or
XX PT more proteins.
XX Claim 1; SEQ ID NO 2461; 134pp; English.
XX PS
XX XX
XX This invention relates to a novel method for altering one or more plant
XX characteristics. Specifically, it refers to identifying genes that are up
XX - or down-regulated in transgenic plants overexpressing the heterodimeric
XX E2Fa/Dpa transcription factor or Arabidopsis and using these sequences to
XX alter plant characteristics accordingly. The present invention describes
XX generating transgenic plants for the production of growth regulators,
XX enzymes, therapeutics, pharmaceuticals and animal feed products, where
XX the altered plant characteristics are selected from increased yield or
XX biomass, enhanced survival capacity, stress tolerance, plant architecture
XX or physiology, altered endoreduplication, biochemistry, signal
XX transduction, storage lipid mobilisation and/or altered photosynthesis,
XX each relative to the corresponding wild type plants. Accordingly, these
XX sequences can also be useful as positive or negative selectable markers
XX during transformation of cells or tissues. The identified genes play a
XX role in a variety of biological processes such as DNA replication, cell
XX wall biosynthesis, nitrogen and/or carbon metabolism or they function as
XX transcription factors. This polynucleotide sequence is thale cress cDNA
XX CC repressed 1.3 fold or more in plants overexpressing the E2Fa/Dpa
XX CC transcription factor, given in an exemplification of the invention.
XX SQ
XX Sequence 954 BP; 288 A; 198 C; 223 G; 233 T; 0 U; 0 Other;
XX
XX Query Match 77.6%; Score 954; DB 12; Length 954;
XX Best Local Similarity 100.0%; Pred.No. 2e-233;
XX Matches 954; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 34 ATGGGACCTCTTAAGTTCATTCATTCCTGCTCTTCTGAGATGCTGAGCAATTG 93
DB 1 ATGGGACCTCTTAAGTTCATTCATTCCTGCTCTTCTGAGATGCTGAGCAATTG 60
QY 94 AGAACCGCTTTTGAAGATGGGGTACGAAGAGACTTGTATCATATCAATCTTGGCTCAC 153
DB 61 AGAACCGCTTTTGAAGATGGGGTACGAAGAGACTTGTATCATATCAATCTTGGCTCAC 120
QY 154 AGAAGTCTGAAACAGAGAAAGTCATCAGGCAAGCATACCAAGAACTTACGGCCGAAGAC 213
DB 121 AGAAGTCTGAAACAGAGAAAGTCATCAGGCAAGCATACCAAGAACTTACGGCCGAAGAC 180
QY 214 CTTCTCAAGACTTTTGAACAGAGACTCTCTCAAGATTTTGAAGAGACTTCTTGTGTGG 273
DB 181 CTTCTCAAGACTTTTGAACAGAGACTCTCTCAAGATTTTGAAGAGACTTCTTGTGTGG 240
```

```
QY 274 ACTTGAACCCGGTGACGGTATGCTTTATTTGGCTAATGAAGCTACAAAAGTGGACT 333
DB 241 ACTTTAAACCCGGTGACGGTATGCTTTATTTGGCTAATGAAGCTACAAAAGTGGACT 300
QY 334 TCAAGCAACCAAGTTCTTATGAGAGTTGCTTGCACAAGACATCAAGCAGCTGCTTAC 393
DB 301 TCAAGCAACCAAGTTCTTATGAGAGTTGCTTGCACAAGACATCAAGCAGCTGCTTAC 360
QY 394 GCTTAGGCAAGCTTACCATGCTCGCTACAGAAGTCTCTTGAAGAGAGCTTGTCTACAC 453
DB 361 GCTTAGGCAAGCTTACCATGCTCGCTACAGAAGTCTCTTGAAGAGAGCTTGTCTACAC 420
QY 454 ACTACCGGTGACTTCAAAAAGCTTTGGTTCTCTTGTACCTCATCAGGTACGAAGA 513
DB 421 ACTACCGGTGACTTCAAAAAGCTTTGGTTCTCTTGTACCTCATCAGGTACGAAGA 480
QY 514 GATGAAGTGAACATGATGATGCTGCTAAGCAAGAGCTAAGCTGCTCATGAGAAATCAAG 573
DB 481 GATGAAGTGAACATGATGATGCTGCTAAGCAAGAGCTAAGCTGCTCATGAGAAATCAAG 540
QY 574 GACAAGCACTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 633
DB 541 GACAAGCACTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
QY 634 ATCAATGCTACTTTTACCGTTTACCAAGATGATGATGATGATGATGATGATGATGATGAT 693
DB 601 ATCAATGCTACTTTTACCGTTTACCAAGATGATGATGATGATGATGATGATGATGATGAT 660
QY 694 GAGGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 753
DB 661 GAGGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
QY 754 ACAAGACCAAGCTTTACTTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 813
DB 721 ACAAGACCAAGCTTTACTTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
QY 814 GATGAAGAGACCTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 873
DB 781 GATGAAGAGACCTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
QY 874 GAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 933
DB 841 GAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
QY 934 CGTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 987
DB 901 CGTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 954
```

RESULT 6  
AB06078  
ID AB06078 standard; DNA; 635 BP.

XX AB06078;

DT 21-AUG-2002 (first entry)

XX Arabidopsis thaliana polynucleotide SEQ ID NO 655.

XX Arabidopsis thaliana; thale cress; plant; transgenic; GMO; disease;  
KM stress; metabolic pathway; biosynthetic pathway; nutrition; fungicide;  
KM insecticide; antibiotic; ds.

XX Arabidopsis thaliana.

XX OS US2002059663-A1.

XX PD 16-MAY-2002.

XX 26-JAN-2001; 2001US-00770149.

XX 27-JAN-2000; 2000US-0178506P.

XX

```
PA (GORL/) GORLACH J.  
PA (ANY/) AN Y.  
PA (HAM/) HAMILTON C M.  
PA (PRIC/) PRICE J L.  
PA (RAIN/) RAINES T M.  
PA (YUY/) YU Y.  
PA (RAME/) RAMEKA J G.  
PA (PAGE/) PAGE A.  
PA (MATH/) MATHEW A V.  
PA (LEDF/) LEDFORD B L.  
PA (WOES/) WOESNER J P.  
PA (HAAS/) HAAS W D.  
PA (GARC/) GARCIA C A.  
PA (KRICK/) KRICKER M.  
PA (SLAT/) SLATER T.  
PA (DAVI/) DAVIS K R.  
PA (ALLE/) ALLEN K.  
PA (HOFF/) HOFFMAN N.  
PA (HURB/) HURBAN P.  
XX  
PI Gorlach J, An Y, Hamilton CM, Price JI, Raines TM, Yu Y,  
PI Rameka JG, Page A, Mathew AV, Ledford BL, Woessner JP, Haas WD,  
PI Garcia CA, Kricker M, Slater T, Davis KR, Allen K, Hoffman N,  
PI Hurban P;  
XX  
DR WPI; 2002-479224/51.  
XX  
PT New nucleic acid that hybridizes to Arabidopsis thaliana sequences,  
PT useful e.g. for preparing transgenic plants with increased resistance or  
PT altered metabolism.  
XX  
PS Claim 1; SEQ ID NO 655; 40pp + Sequence Listing; English.  
XX  
CC The invention relates to nucleic acids (i) that hybridise under stringent  
CC conditions to any of 999 sequences (AB065424-AB066422) or their  
CC fragments, (ii) are used to express the corresponding polypeptides (ii) or  
CC to produce genetically modified plant cells or transgenic plants, which  
CC may have improved resistance to disease or stress, or altered  
CC metabolic/biosynthetic pathways (for production of commercial,  
CC nutritional or medicinal products), or generally any trait of interest,  
CC or can be used to screen for biologically active agents (e.g. fungicides,  
CC insecticides and antibiotics). Note: The sequence data for this patent  
CC did not form part of the printed specification, but was obtained in  
CC electronic format directly from the USPTO at  
CC seqdata.uspto.gov/sequence.html?docID=99909770149  
XX  
SQ Sequence 635 BP; 197 A; 133 C; 147 G; 158 T; 0 U; 0 Other;
```

Query Match 51.6%; Score 635; DB 6; Length 635;  
Best Local Similarity 100.0%; Pred. No. 5.8e-152;  
Matches 635; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 16 ACTTAAAGTAGAAGAAAATGGGCACTTTAAGTTTGTGATTTCTGCTCGTCTTCT 75
DB 1 ACTTAAAGTAGAAGAAAATGGGCACTTTAAGTTTGTGATTTCTGCTCGTCTTCT 60
QY 76 GATGATGCTGAGCAATTTGAAGACCGCTTTTGAAGATGGGGTACGAAGAGACTTGAATC 135
DB 61 GATGATGCTGAGCAATTTGAAGACCGCTTTTGAAGATGGGGTACGAAGAGACTTGAATC 120
QY 136 ATATCAATCTTGGCTACAGAAAGTGTGAACAGAGAAAGTCAATCAGGCAAGTATCCAC 195
DB 121 ATATCAATCTTGGCTACAGAAAGTGTGAACAGAGAAAGTCAATCAGGCAAGTATCCAC 180
QY 196 GAAACCTACGCGGAGAACCTTCTCAAGACTTTTGAAGAGCTCTTAACGATTTGAG 255
DB 181 GAAACCTACGCGGAGAACCTTCTCAAGACTTTTGAAGAGCTCTTAACGATTTGAG 240
QY 256 AGAGCTATCTTGTGTGAATCTTGAACCCGGTGAAGCGTATGCTTATTTGGCTAAATGA 315
DB 241 AGAGCTATCTTGTGTGAATCTTGAACCCGGTGAAGCGTATGCTTATTTGGCTAAATGA 300
QY 316 GCTTACAAAAGATGAGCTTCAAGCAACAAAGTTCTTATGAAAGTCTTGCAACAGACA 375
```



Accession	Sequence	Position
D6	301 GCTACAAAAGATGGACTTCAAGCAACCAAGTTCTTATGAAAGTGTGCAACAAGACA	360
Qy	376 TCAACGAGCTGCTTACGCTAGGACAGCTTACATGCTGCTCAAGAAAGTCTTTGAA	435
D6	361 TCAACGAGCTGCTTACGCTAGGACAGCTTACATGCTGCTCAAGAAAGTCTTTGAA	420
Qy	436 GAGGACGTTGCTCAACAACACTACCGGTGACTTCAGAAAGCTTTGGTTTCTCTTGTAAC	495
D6	421 GAGGACCTTGCTTACCACTACCGGTGACTTCAGAAAGCTTTGGTTTCTCTTGTAAC	480
Qy	496 TCATACAGGTGACGAAGGAGTGAAGTGAACATGATGGCGTAAAGCAAGAGCTTAAGCTG	555
D6	481 TCATACAGGTGACGAAGGAGTGAAGTGAACATGATGGCGTAAAGCAAGAGCTTAAGCTG	540
Qy	556 GTCCATAGAAATTCAAAGCAACGACTACATGATGAGATGTTATTAGATCTTGTC	615
D6	541 GTCCATAGAAATTCAAAGCAACGACTACATGATGAGATGTTATTAGATCTTGTC	600
Qy	616 ACAAGAAGCAAGCTCAGTCAATGCTACTTTTAA	650
D6	601 ACAAGAAGCAAGCTCAGTCAATGCTACTTTTAA	635

RESULT 7	
ADR63461	
ID ADR63461	standard; cDNA; 1184 BP.

DT	02-DEC-2004 (first entry)
XX	
DE	Cotton cDNA sequence, SEQ ID 4242

KM Cotton; ss; plant; cold tolerance; growth rate; cell cycle pathway;  
KM drought tolerance; plant disease resistance; galactosaminan; lignin;  
KM plant growth regulator; heat tolerance; herbicide tolerance;  
KM homologous recombination; extreme osmotic condition tolerance;  
KM stress resistance; pest resistance; yield; photosynthesis; seed oil;  
KM stress resistance.

OS Gossypium hirsutum.  
XX  
PN US2004181830-A1.

PD	16-SEP-2004	
XX		
PF	29-JAN-2004	2004US-00767795.

PA (KOVÁ/) KOVALIC D K.  
PA (ZHOU/) ZHOU Y.  
PA (CAOY/) CAO Y.

PI Kovalic DK, Zhou Y, Cao Y;  
XX  
DR WPI; 2004-667718/65.

PT New recombinant nucleic acid molecules and polypeptides from *Gossypium hirsutum*, useful for producing plants with improved biological characteristics (e.g. improved plant cold or drought tolerance).

PS Claim 1; SEQ ID NO 4242; 14pp; English.

CC The invention relates to a recombinant polynucleotide comprising any of  
CC the 58798 cotton plant cDNA sequences mentioned in the specification.  
CC Also a recombinant polypeptide comprising any of the 58798 amino acid  
CC sequences mentioned in the specification and producing a plant having an  
CC improved property. Producing a plant having an improved property  
CC comprises transforming a plant with a recombinant construct comprising a  
CC promoter region functional in a plant cell operably joined to a

polynucleotide comprising a coding sequence for a polypeptide associated with the property, and growing the transformed plant. The polypeptide is useful for improving plant cold tolerance, manipulating growth rate in plant cells by modification of the cell cycle pathway, improving plant drought tolerance, providing increased resistance to plant disease, producing galactosemannan (or lignin or plant growth regulators), increasing plant heat tolerance, improving plant tolerance to herbicides, increasing the rate of homologous recombination in plants, improving plant tolerance to extreme osmotic conditions or to pathogens or pests, improving yield by modification of photosynthesis, modifying seed oil or protein yield and/or content, improving uptake, or modification of carbohydrate, nitrogen or phosphorus use and/or uptake, or improving yield by providing improved plant growth and development under at least one stress condition. The polynucleotide and polypeptide may also be used in recombinant DNA constructs, in physical arrays of molecules, as plant breeding markers, or in computer-based storage and analysis systems. The present sequence is a cotton plant cDNA of the invention. NOTE: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from USPO at [seqlata.uspto.gov/sequence.html?DOCID=20040181930](http://seqlata.uspto.gov/sequence.html?DOCID=20040181930). However only 6565 polynucleotide sequences were available, the remaining 5213 polynucleotides and all 58798 protein sequences were not present.

Sequence 1184 BP; 375 A; 233 C; 281 G; 295 T; 0 U; 0 Other;

Query Match	40.5%;	Score 497.8;	DB 13;	Length 1184;
Best Local Similarity	70.7%;	Pred. No. 8:5e-117;		
Matches 691; Conservative	0;	Mismatches 282;	Indels 4;	Gaps 2

DY  
DB

	12	AACACTAAAGTAGAAGA-AAATGGGCACTTAAAGTTTCGATTCTGTCCCTGC	70
	47	AAGAATAAAGGAAGAGACATGGCACCTTACAGTGCCACGACGACTTCCTCGG	10

Db 347 ATGAAGCACCAAAAGGTGAACTCAAGCAATCAGTCTTATGGAATAGCTTCACAA 40

QY 311 ATGAAGCTACAAAAGATGAGCTTCAACGACCAAGTCTTATGGAAGTTGCTGCACAA 37  
Db 347 ATGAAGCCACCAAAAGGTGAGCTTCAAGCAATCAGGCTCTATGGAATAAGCTGCACAA 40

Db 167 TAATCATAGATATATTCGGGTCAACAGAAATGCCAGCAACGAACCTTGATTCCGAAAACCT 22

Query Match	40.5%;	Score 497.8;	DB 13;	Length 1184;
Best Local Similarity	70.7%;	Pred. No. 8:5e-117;		

Qy 671 GCGAGAAATTCCTCAAGAGCTTGGAGAGAGATGATGATGACAGTTCCTTGCACCTTT 730  
 Db 707 GAATGACATTAACAGAGACTT---GAAGCTGATCTTAAGATGATGATTCCTTGCACCTTAC 763  
 Qy 731 TGAAGTCAACCATTCAGTGTCTTGAACAAGCAGACTTACTTTGTGATGATTCCTTGCCTT 790  
 Db 764 TAAGTCCACAGTGAAGTGTGTGCTATCTCCGAAAAGTATTTTGAAGAGTTCCTTGCC 823  
 Qy 791 CAGAAATCAACAAACTGGAATGATGATGAGAGACATCTAGATTTGTGACCAAGAG 850  
 Db 824 TAGCAATCAATAGACGAGAAACGATAGAGAGCTCTTACTTATGATGATTCCTTGCCTT 883  
 Qy 851 CTGAGATTCATTCAGAGTCAATTCGAGAGAGTACACGCGAGAGACAGATTCCTTGCCTT 910  
 Db 884 CTGAGATTCATTCAGAGTCAATTCGAGAGAGTACACGCGAGAGACAGATTCCTTGCCTT 943  
 Qy 911 AGAAAGTATTCACCAAGACCTGCGAGATTCAGAGAGATTCGCGACCTTTCG 970  
 Db 944 CTGCGTCAATTCAGAGACCTGAGAGATTCAGAGAGATTCGCGACCTTTCG 1003  
 Qy 971 GTGAGATGATGCTTAA 987  
 Db 1004 GACATGTGAGAAATTTGA 1020

## RESULT 8

ADB94854

ID ADB94854 standard; cDNA; 1058 BP.

XX AC ADB94854;

XX DT 04-DEC-2003 (first entry)

XX DE Programmed cell death pathway protein annexin cDNA #6.

XX programmed cell death; plant development; plant cell cycle; ATL2; DAD1;  
 KW Dnaase; 11s; 1sdt; nucellin-like aspartic protease; annexin; prohibitin;  
 KW fan-like protein; race; retinoblastoma-related protein; SIVA; TRFD;  
 KW TBGT; xylogenic kinase; put-alpha; cyteine protease; RPPs-like protein;  
 KW gp 91 NADPH oxidase subunit; NPR-like protein; BAG-1;  
 KW defender against cell death; lethal leaf spot; lesion stimulating death;  
 KW seven in absentia; transcription initiation factor;  
 KW testis enhanced gene transcript; gene; ss.

XX OS Eucalyptus grandis.

XX PN US2003082724-A1.

XX PD 01-MAY-2003.

XX PF 14-AUG-2002; 2002US-00219220.

XX PR 04-JUN-1999; 99US-00325932.

XX PA (GENE-) GENESIS RES &amp; DEV CORP LTD.

XX PI Flim B, Lasham A;

XX DR MPI; 2003-786916/74.

XX DR P-PSDB; ADB94896.

XX PT New isolated polynucleotide useful for modulating programmed cell death,  
 PT altering the development cycle of plant cells, and subsequently modifying  
 PT plant development.

XX PS Claim 1; Page 138; 214pp; English.

XX CC The invention describes an isolated polynucleotide (I) comprising a  
 CC nucleotide sequence that is one of 145 fully defined sequences of 221-  
 CC 315 base pairs (bp), given in the specification, and/or its complements,  
 CC reverse complements, reverse sequences, or sequences having 75, 90 or 95  
 CC % sequence identity to or that hybridise under stringent hybridisation  
 CC conditions to one of the 145 sequences. The methods and compositions of

CC the present invention to do with (I) are useful for modulating programmed  
 CC cell death and thereby altering the development cycle of plant cells, and  
 CC altering plant development. This sequence encodes a protein associated  
 CC with the programmed cell death pathway.

XX Sequence 1058 BP; 306 A; 227 C; 274 G; 251 T; 0 U; 0 Other;

Query Match 38.5%; Score 473.4; DB 10; Length 1058;

Best Local Similarity 69.0%; Pred. No. 1,4e-110;

Matches 664; Conservative 0; Mismatches 296; Indels 3; Gaps 1;

Qy 25 AGAAGAAAATGCGCATCTTAAAGTTCCTGATTCCTGCTCTTCTGATGATGCT 84  
 Db 63 AGAAGAGATGCGCATCTGCGGTGCGACCTCGCTCCGTCCGCTGAGATGCC 122  
 Qy 85 GAGCAATGGAACCGCTTTTGAAGATGCGGTAGCAAGAGAGACTTATATCATC 144  
 Db 123 GAGCAGCTCCAAAAGCTTTGCGAGATGGGAGAACAAATGATGATCATATCAT 182  
 Qy 145 TTGGCTCACGAAGTGTGTAACAGAGAAAGTATATGAGCAAGATACCAAGAACTTAC 204  
 Db 183 CTGCTCACAGAAACGAGCGCAGCGAAAGTATCCGAACATATGCCGAGATAT 242  
 Qy 205 GCGAAGACCTTCTCAAGACTCTTGAACAAGAGCTCTTAAGATTTCCAGAGAGCTATC 264  
 Db 243 GGGAAAGATCTTTCAAGCGCTTGAACAAGAACTCTTAAGATTTGAGATTTCTG 302  
 Qy 265 TTGTTGTGACCTTTGAACCCGCTGAGCGTATGCTTTATTTGCTATGAGCTACAAA 324  
 Db 303 CTTCGTGAGACCTTGATCTGCGAGCGGTATGATCTTGTCCATGATGATGATCAAG 362  
 Qy 325 AGATGACTTCAAGCAACCAAGTCTTATGAAAGTTCCTTGCAAGAGATCAAGCGAG 384  
 Db 363 AGATTAATCTTCAAGCAACCTGAGTCTTCAATGAAATGCTTGACGAGTCTTCAATGAG 422  
 Qy 385 CTGCTTCAAGCTGAGAGCTTACATGCTGCTGCTTACAGAAAGCTCTTGAAGAGCGTT 444  
 Db 423 TTATTCATGTGAGGAGCGCTATACGCTCGTTATTAAGAAATCTCTTGAAGAAATC 482  
 Qy 445 GCTCACACACATACCGGTATCTTCAAGAAAGCTTTGCTTCTTGTATCTTATCAAGG 504  
 Db 483 GCATATCACATCTATCGGGATTTCCGCAAGCTGCTTCTCTGCAAGTATCTTTCGG 542  
 Qy 505 TACGAAGAGATGAAATGAAATGATGATGCTTGAAGCAAGAGCTTAACTGCTATGAG 564  
 Db 543 TATGAGGGGCTGAGGTGAAATGATGATGCTGAGGCTAAGATTAATCTTCAATGAG 602  
 Qy 565 AAAATCAAGAGCAAGCACTAATGATGATGATGATGATGATGATGATGATGATGATG 624  
 Db 603 AAGATTCAGAGAGGCTTACATGATGATGATGATGATGATGATGATGATGATGATG 662  
 Qy 625 AAGCTCAGATCAATCTTCTTAAACGCTTACCAAGATGATGATGATGATGATGATGAT 684  
 Db 663 AAGCTCAGCTTAAATCAACCTCAATTAATCAACAAAGATGATGATGATGATGATGAT 722  
 Qy 685 AAGATCTTGAAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 744  
 Db 723 AAGATCT---GAAGCTGATCAAAATGATGATGATGATGATGATGATGATGATGAT 779  
 Qy 745 CAGTCTTCAAGAGCAAGCTTATGATGATGATGATGATGATGATGATGATGATGATG 804  
 Db 780 AAGTCTTCACTTATCAAGAGATGATGATGATGATGATGATGATGATGATGATGATG 839  
 Qy 805 ACTGAACTGATGAGAGAGAGCTCACTAGATTTGTGACCAAGAGCTGATGATGATGAT 864  
 Db 840 CTGGAAACAGATGAATGGGCTTTACAGAGATGATGATGATGATGATGATGATGATG 899  
 Qy 865 AAGTATTTGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 924  
 Db 900 CAGCGATTAAGAAACAGTACCAAAAGAAACAGTGTCTCTGATGATGATGATGATGAT 959  
 Qy 925 AAAGACATCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 984

Db 960 AATGACATCTGGGACTACAGAAGATCTTCGCTTTGGTGGACATGAGATGCT 1019  
QY 985 TAA 987  
Db 1020 TGA 1022

RESULT 9  
AAA67100  
ID AAA67100 standard; DNA; 2588 BP.  
XX  
AC AAA67100;  
XX  
DT 31-OCT-2000 (first entry)  
XX  
DE Eucalyptus grandis annexin nucleotide sequence SEQ ID NO:28.  
XX  
KM Eucalyptus grandis; pinus radiata; Monterey pine; modification;  
KM plant cell wall; polysaccharide; polysaccharide biosynthetic pathway;  
KM transgenic plant; ds.  
XX  
OS Eucalyptus grandis.  
XX  
PN WO200022092-A2.  
XX  
PD 20-APR-2000.  
XX  
PF 08-OCT-1999; 99WO-NZ000169.  
XX  
PR 13-OCT-1998; 98US-00170862.  
PR 11-AUG-1999; 99US-0148426P.  
XX  
PA (GENE-) GENESIS RES & DEV CORP LTD.  
PA (FLET-) FLETCHER CHALLENGE FORESTS LTD.  
XX  
PI Blokeberg LN;  
XX  
DR WPI; 2000-339328/29.  
DR P-PSDB; AAB16294.  
XX  
PT New genes encoding proteins involved in a plant polysaccharide  
PT biosynthetic pathway, useful for modulating or altering the  
PT polysaccharide content, composition or structure of the plant.  
XX  
PS Claim 1; Page 48-49; 301pp; English.  
XX  
CC The present invention describes isolated polynucleotides (PN) comprising  
CC a sequence selected from one of 835 nucleotide sequences given in  
CC AAA67073 to AAA67907, their (reverse) complements, sequences producing an  
CC expectation (E) value of 0.01 or less compared to the 835 sequences,  
CC sequences at least 50% identical to them, 200, 100, 40 or 20-mers of the  
CC 835 sequences or sequences that are degenerately equivalent or allelic to  
CC the 835 sequences. The polynucleotides are used to modify the activity of  
CC a polypeptide involved in a polysaccharide biosynthetic pathway in the  
CC plant. They are especially used to modulate or alter the polysaccharide  
CC content, composition or structure of the plant. AAB16268 to AAB16340 are  
CC proteins encoded by some of the polynucleotide sequence given in the  
CC present invention  
XX  
SQ Sequence 2588 BP; 715 A; 579 C; 628 G; 666 T; 0 U; 0 Other;

Query Match 38.4%; Score 471.8; DB 3; Length 2588;  
Best Local Similarity 68.6%; Pred. No. 5e-110;  
Matches 666; Conservative 0; Mismatches 302; Indels 3; Gaps 1;

QY 30 AAAAATGGCGACTCTTAAGGTTTCTGATTTCTGCTCTCTTGATGATGCTGAGCA 89  
XX  
DB 1386 AACAAATGTGACTCTGACCGTCCGACGACCTGCCCTGTAGCGGATGACTGCGAGCA 1445

QY 90 ATTGAGAACCGCTTTTGAAGAGTGGTACGACGAGATTTGATATATATCATCTTGGC 149  
XX  
DB 1446 GCTCCGACAGCCCTTCGACGATGGGACAAATAGAGAACTATCATATCATATTGGG 1505

QY 150 TCACAGAAAGTGTGAAACAGAGAGAAATGATCATGCGAAGATATCCAGAAACCTACGCGCA 209  
|||  
DB 1506 TCATAGGATATGCGGGCGAGAGAGAACTGATTTGGCAAACTATGCGAGACTTACGGCGCA 1565  
QY 210 AAGCTTTCAAGACTCTTGAACAGAGAGTCTCTAAGATTTTGGAGAGCTATCTTGT 269  
|||  
DB 1566 GGACCTCCTCAAGGACATTGAGCAGAGAACTTACCAATGATTTTCGAGAGGCTGGTGGTCT 1625  
QY 270 GTGGACTCTGAACCGGAGAGGAGCTTATTTGGCTAATGGAAGCTACAAAAGATG 329  
|||  
DB 1626 TTGGTACTTGAATCCGGCTGAACGTGATGCGTACTTTGGCAATGAGACGAAAGATG 1685  
QY 330 GACTTCAGCAACCAAGTCTTATGGAAGTTCCTGACCAAGACATCAACCGACCTGCT 389  
|||  
DB 1686 GACTTCAGCAACCAAGTCTTATGGAAGTTCCTGACCAAGACATCAACCGACATGCTGCT 1745  
QY 390 TCACGCTGAGGCAAGCTTACATGCTGCTACAGAAAGTCTCTTGAAGAGAGAGTGTCTCA 449  
|||  
DB 1746 TATGGCAAGACAAAGATATCATGCCGATACAAAGAGTCTGGAAGAGAGAGTGGTCA 1805  
QY 450 CCACACTACCGGAGCTTCAGAAAGCTTTGGTTCTCTGTACTCATACAGGATGCA 509  
|||  
DB 1806 CCAACACTGAGATTTTCGTAAGTTGCTGTAACCTCTTGAGCTCTTACCGTTATGA 1865  
QY 510 AGGAGTGAAGTGAACATGACATGAGCTTAAGCAAGAGCTAAGCTGTCCATGAGAAAT 569  
|||  
DB 1866 TGAAGTGAAGTGAATATGACTTTGGCAAAAGAGGCTAAGATATCTTCAAGAAAGAT 1925  
QY 570 CAAGCAAGCACTACATGATGAGATGTTTGAATCTTGTCCACAAAGCAAAAGC 629  
|||  
DB 1926 CTCAGAGAGGCTTATGGCATGAGAGATCTCTAAGGATTTTGGTACTACGAGCAAAAGC 1985  
QY 630 TCAGATCAATGCTACTTTTAAACGTTAACCAAGATGATCATGCGAGAAATCTCAAGAG 689  
|||  
DB 1986 AAGGTCATATGCTTACGCTGAATCACTACAAAATAGTTGAAATGATATCAACAAGA 2045  
QY 690 TCTTGAAGAGAGATGATGATGACAAAGTCTCTGCACTTTTGAAGTCAACCATTCAGTG 749  
|||  
DB 2046 TTT--GAAAACTGATCCAAAAGAGCGGTTCTTACTATATGAGAGCTACAGTAAAGTG 2102  
QY 750 CTTGACAAAGCCAGAGCTTTACTTTGTGATGTTCTTGTGATGCAATTAACAAACTGG 809  
|||  
DB 2103 CCTGACTCCCTGAGAAAGTATTTGAAAAGGTTCTTGTCTACGATCAATTAACGAGG 2162  
QY 810 AACTGATGAAGAGCACTGACTGATGATGATGATGATGATGATGATGATGATGATGAT 869  
|||  
DB 2163 AACAGATGAAGGGGCTCTGACCAAGTATGTTCTACAGGCGGAGTTGACATGAATG 2222  
QY 870 CATTGAGAGAGATACCAAGCGAGGAAACGATTCCTTTGAGAAAGCTATTATCCAAAGA 929  
|||  
DB 2223 TATTAAGTGAAGATACCAAGAGGAAATGATCCCTCTCGATCGTGCATTTGTCAAGGA 2282  
QY 930 CACTGTGAGATTTACGAGAAATGCTGTGCCACTTCTCGGTGAAGATGATGCTTAATC 989  
|||  
DB 2283 CACTGTGAGACTATGAAAAATGCTTGTGCATTGATTTGCGACAGTGCAGGCTTGAT 2342  
QY 990 AATCAATCCTC 1000  
|||  
DB 2343 TACAGTACTC 2353

RESULT 10  
ADE8194/c  
ID ADE8194 standard; cDNA; 487 BP.  
XX  
AC ADE8194;  
XX  
DT 29-JAN-2004 (first entry)  
XX  
DE Arabidopsis thaliana expressed polynucleotide seq id 765.  
XX  
KM genetically modified organism; transgenic organism; plant;  
KM inhibitor testing; activator testing; modifier testing; fungicide;

KM insecticide; genetic function; genetic regulation; cellular metabolism;  
 KM gene; ss.  
 XX Arabidopsis thaliana.  
 OS  
 PN US2003115639-A1.  
 XX  
 PD 19-JUN-2003.  
 XX  
 PF 26-JAN-2001; 2001US-00770961.  
 XX  
 PR 27-JAN-2000; 2000US-0178466P.  
 XX  
 PA (GORL/) GORLACH J.  
 PA (ANY/) AN Y.  
 PA (HAM/) HAMILTON C M.  
 PA (PRIC/) PRICE J L.  
 PA (RAIN/) RAINES T M.  
 PA (YUY/) YU Y.  
 PA (RAME/) RAMEKA J G.  
 PA (PAGE/) PAGE A.  
 PA (MATH/) MATHEW A V.  
 PA (LED/) LEDFORD B L.  
 PA (WOES/) WOESSNER J P.  
 PA (HAAS/) HAAS W D.  
 PA (GARC/) GARCIA C A.  
 PA (KRIC/) KRICKER M.  
 PA (SLAT/) SLATER T.  
 PA (DAVI/) DAVIS K R.  
 PA (ALLE/) ALLEN K.  
 PA (HOFF/) HOFFMAN N.  
 PA (HURB/) HURBAN P.  
 XX  
 PI Gorlach J, An Y, Hamilton CM, Price JL, Raines TM, Yu Y,  
 PI Rameka JG, Page A, Mathew AV, Ledford BL, Woessner JP, Haas WD;  
 PI Garcia CA, Kricker M, Slater T, Davis KR, Allen K, Hoffman N;  
 PI Hurban P;  
 PI  
 XX  
 DR WPI; 2003-810930/76.  
 XX  
 PT Novel Arabidopsis thaliana nucleic acids useful for generating  
 PT genetically modified transgenic organisms, for screening biologically  
 PT active agents such as fungicides, insecticides.  
 XX  
 PS Claim 1; SEQ ID NO 765; 44bp; English.  
 XX  
 CC The invention describes a nucleic acid (I) comprising a sequence capable  
 CC of hybridizing under stringent conditions to any one of 999 fully defined  
 CC Arabidopsis thaliana sequences (1) as given in specification e.g., 360,  
 CC 1137, 455, 219, 472 nucleotides, etc, or its fragment. (I) is useful as a  
 CC hybridization probe to complementary molecules in a cDNA library. (I) is  
 CC also useful for generating genetically modified and transgenic organisms,  
 CC usually plant cells and plants. A protein encoded by (I) is useful in  
 CC screening assays to determine the effect of candidate inhibitors,  
 CC activators or modifiers of the gene product. The protein is also useful  
 CC for screening biologically active agents e.g., fungicides and  
 CC insecticides. A genetically modified cell, comprising an exogenous  
 CC nucleic acid, where the nucleic acid comprises transcription regulatory  
 CC sequences operably linked to a sequence capable of hybridizing under  
 CC stringent conditions to (1) is useful in the study of genetic function  
 CC and regulation, for alteration of the cellular metabolism and for  
 CC screening compounds that may affect the biological function of the gene  
 CC or gene product. This sequence represents an Arabidopsis thaliana  
 CC polynucleotide of the invention.  
 XX  
 SO Sequence 487 BP; 136 A; 96 C; 88 G; 167 T; 0 U; 0 Other;  
 Query Match 38.2%; Score 470; DB 10; Length 487;  
 Best Local Similarity 100.0%; Pred.No. 7.8e-110;  
 Matches 470; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 761 CAGAGCTTACTTGTGATGTTCTTGTGATGATCAAAACTGGAAGTATGAG 820  
 |||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 487 CAGAGCTTACTTGTGATGTTCTTGTGATGATCAAAACTGGAAGTATGAG 428  
 QY 821 GAGCACTCACTAAGATTGTGACCAAGAGCTGAGATTGACTTGAAGTCAATTGAGAGG 880  
 |||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 427 GAGCACTCACTAAGATTGTGACCAAGAGCTGAGATTGACTTGAAGTCAATTGAGAGG 368  
 QY 881 AGTACAGGCGCAGAGAACAGATTCTTTGGAGAAAGCTATTACCAAGACACTCGTGGAG 940  
 |||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 367 AGTACAGGCGCAGAGAACAGATTCTTTGGAGAAAGCTATTACCAAGACACTCGTGGAG 308  
 QY 941 ATTACAGAGAGATGCTGCTGACACTTCTGAGTGAAGATGATGCTTAATCAATCAATCTTC 1000  
 |||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 307 ATTACAGAGAGATGCTGCTGACACTTCTGAGTGAAGATGATGCTTAATCAATCAATCTTC 248  
 QY 1001 CACAGAGAAACATAGCTGCTGACAGCTTCTGTATCTCTTATCTTCTCTCTCTCTCT 1060  
 |||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 247 CACAGAGAAACATAGCTGCTGACAGCTTCTGTATCTCTTATCTTCTCTCTCTCTCTCT 188  
 QY 1061 TTGATGATGTTGCAAACTGTTGATTTTGTCTTCAAAAAACCTTGTGTTCTGTTGTG 1120  
 |||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 187 TTGATGATGTTGCAAACTGTTGATTTTGTCTTCAAAAAACCTTGTGTTCTGTTGTG 128  
 QY 1121 TGTGTTGATGTTCTTAATAATGCAAAAGAGAGACAGAGAACCAAGTGTGCTCTTCA 1180  
 |||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 127 TGTGTTGATGTTCTTAATAATGCAAAAGAGAGACAGAGAACCAAGTGTGCTCTTCA 68  
 QY 1181 AGTTATATATATATATGAGAGCATTTGGCTTAAAAAATTTTTTTTTTTTTTTTTT 1230  
 |||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 67 AGTTATATATATATATGAGAGCATTTGGCTTAAAAAATTTTTTTTTTTTTTTTTT 18  
 RESULT 11  
 ID AAA67128 standard; DNA; 1293 BP.  
 XX  
 AC AAA67128;  
 XX  
 DT 31-OCT-2000 (first entry)  
 XX  
 DE Pinus radiata annexin nucleotide sequence SEQ ID NO:109.  
 XX  
 KM Eucalyptus grandis; pinus radiata; Monterey pine; modification;  
 KM plant cell wall; polysaccharide; polysaccharide biosynthetic pathway;  
 KM transgenic plant; ds.  
 XX  
 OS Pinus radiata.  
 XX  
 PN WC200022092-A2.  
 XX  
 PD 20-APR-2000.  
 XX  
 PF 08-OCT-1999; 99WC-NZ000169.  
 XX  
 PR 13-OCT-1998; 98US-00170862.  
 PR 11-AUG-1999; 99US-0148426P.  
 XX  
 PA (GENE-) GENESIS RES & DEV CORP LTD.  
 PA (FLEET-) FLETCHEER CHALLENGE FORESTS LTD.  
 XX  
 PI Blokeberg LN;  
 XX  
 DR WPI; 2000-339328/29.  
 DR P-PDB; AAB16321.  
 XX  
 PT New genes encoding proteins involved in a plant polysaccharide  
 PT biosynthetic pathway, useful for modulating or altering the  
 PT polysaccharide content, composition or structure of the plant.  
 XX  
 PS Claim 1; Page 93; 301bp; English.  
 XX  
 CC The present invention describes isolated polynucleotides (PN) comprising  
 CC a sequence selected from one of 835 nucleotide sequences given in  
 CC AAA67073 to AAA67907, their (reverse) complements, sequences producing an

CC Expectation (E) value of 0.01 or less compared to the 835 sequences,  
 CC sequences at least 50% identical to them. 200, 100, 40 or 20-mers of the  
 CC 835 sequences or sequences that are degenerately equivalent or allelic to  
 CC the 835 sequences. The polynucleotides are used to modify the activity of  
 CC a polypeptide involved in a polysaccharide biosynthetic pathway in the  
 CC plant. They are especially used to modulate or alter the polysaccharide  
 CC content, composition or structure of the plant. AAB16268 to AAB16340 are  
 CC proteins encoded by some of the polynucleotide sequence given in the  
 CC present invention

XX Sequence 1293 BP; 359 A; 290 C; 313 G; 331 T; 0 U; 0 Other;

Query Match 38.1%; Score 468.6; DB 3; Length 1293;

Best Local Similarity 68.4%; Pred. No. 2.5e-109;

Matches 664; Conservative 0; Mismatches 304; Indels 3; Gaps 1;

QY 30 AAAAATGGGCGACTCTTAAGGTTTCTGATTCCTGCTCTTCTGATGATGCTGAGCA 89  
 DB 91 AACAAATGTCGACTCTGACCGTCCCGACGCACTGCCCTGTGACCGGATGACTGCGAGCA 150  
 QY 90 ATTGAGAACCGCTTTGAAGATGGGGTACGACGAGGACTTGATGATATCAATCTTGGC 149  
 DB 151 GCTCCGAGACGCTTCGCGAGATGGGAAACAATGAAAGCTGATCATATTCATTTGGG 210  
 QY 150 TCACAGAAAGTGTGAACAGAGAAAGTCATCAGGCAAGCATACCAAGAACTTACGGCGA 209  
 DB 211 TCATAGGAATGGGGCGGAGAGAGAGCTGATTGGCAAACTATGCGGAGCTTACGGCGA 270  
 QY 210 AGACCTTCTCAGACTCTTGACAGAGAGCTCTTAAAGATTTGAGAGAGCTATCTTGT 269  
 DB 271 GGACCTCTCAGGCAATGAGACAGAGAACTTACCAATGATTTGAGAGGCTGTGTCTCT 330  
 QY 270 GTGAGCTCTTGAACCGGAGGAGGTGATGCTTATTTGGCTAAAGAGCTACAAAGAGT 329  
 DB 331 TTGCTCATCTTGATCCGCTGCTAAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 390  
 QY 330 GACTTCAAGCAACCAAGTCTTATGGAAGTTGCTTGACAGAGCATCAACGAGCTGCT 389  
 DB 391 GACTTCAAGCAACCAAGTCTTATGGAAGTTGCTTGACAGAGCATCAACGAGCTGCT 450  
 QY 390 TCACGCTAGGCAAGCTTACATGCTGCTACAGAAAGTCTCTTGAAGAGGACTTGGCTCA 449  
 DB 451 TATGGCAAGCAAGCATATCATGCCGATCAAGAAAGTCAATGAAAGAGGACCTCCCTCA 510  
 QY 450 CCACTACCGGAGCTTGCAAGAAAGTCTTGTGCTTGTGCTGCTGCTGCTGCTGCTGCTGCT 509  
 DB 511 CCACTACCGGAGCTTGCAAGAAAGTCTTGTGCTTGTGCTGCTGCTGCTGCTGCTGCTGCT 570  
 QY 510 AGGAGATGAGTGAATGATGATCATTTGGCTAAGCAAGAGCTAGCTGCTCATGAGAAAT 569  
 DB 571 TGGAGATGAGTGAATGATGATCATTTGGCAAAAGCAAGAGCTAGATCTCCAGAGAAAGT 630  
 QY 570 CAAGGACAGCACTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 629  
 DB 631 CTCAGGAAGGCTTATGCGCATAGAGATCTCATAGAGATTTTGGCTACTATGAGCAAGC 690  
 QY 630 TCAGATCAATGCTACTTTTAAACGTTAAGCAAGATGATGATGATGATGATGATGATGATG 689  
 DB 691 ACAGGTCATGCTACGCTGATATCACTACAAAAATGATTTGGAATGATATCAACAAAGGA 750  
 QY 690 TCTTGAAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 749  
 DB 751 TTT---GAAAACGATCCAAAAGACGGTTCCTTACTATATGAGAGCTACATTAAGT 807  
 QY 750 CTTGACAGAGCAGAGCTTACTTTGTGATGATTTCTTGTTCAGCAATCAAAAATGG 809  
 DB 808 CTTGACAGAGCAGAGCTTACTTTGTGATGATTTCTTGTTCAGCAATCAAAAATGG 867  
 QY 810 AACTGATGAGAGGAGCTCACTAGAAATTTGACCAAGAGCTGATGATGATGATGATGATGATG 869  
 DB 868 AACGATGAGAGGAGCTCTGACAGAGATGTTGCTACACAGGCGGAGGTTGACATGAAAGT 927  
 QY 870 CATTTGAGAGAGTACACAGCGCAGAGAACAGATTCCTTTGGAAGAGCTTATCAAGA 929

DB 928 TATTAAGTGAAGAGTACACAGAGGAAATGATCTCCTTCATGCGCATGTCAGAGA 987  
 QY 930 CACTGTGAGATTTACAGAAATGCTGCTGCCACTTCTCGGTGAAGATGATGATGATGATG 989  
 DB 988 CACTACTGAGACTATGAAAAATGCTTGTGCAATGATTTGGCCACGTCGAGGCTTGATT 1047  
 QY 990 AATCAATCCTC 1000  
 DB 1048 TACAGTACTC 1058

RESULT 12

AAFA4756

ID AAFA4756 standard; cDNA; 1293 BP.

XX AAFA4756;

XX 27-MAR-2001 (first entry)

XX Annexin-like protein coding sequence #1.

XX Cell death modulator; programmed cell death; PCD; apoptosis;

XX forestry plant; ss.

XX Pinus radiata.

XX MO200075331-A1.

XX 14-DEC-2000.

XX 02-JUN-2000; 2000WO-NZ000086.

XX 04-JUN-1999; 99US-00325932.

XX (GENE-) GENESIS RES & DEV CORP LTD.

XX (FLEET-) FLETCHEER CHALLENGE FORESTS LTD.

XX Flinn B, Lasham A;

XX WPI; 2001-061724/07.

XX P-PSDB; AAB65730.

XX Novel defender against cell death polynucleotide useful for modulating

XX programmed cell death pathway and specific development pathways in

XX forestry plant.

XX Claim 1; Page 54-55; 142pp; English.

XX The present invention relates to coding sequences (see AAF44740-F44840

XX and AAF44843-F44844) and proteins (see AAB65714-B65814) involved in

XX programmed cell death (PCD; apoptosis). The coding sequences and proteins

XX of the present invention are useful for modulating a PCD or cell death

XX pathway and various developmental pathways in a forestry plant, by stably

XX incorporating one of the present coding sequences into the genome of a

XX forestry plant, where the coding sequence provides a PCD pathway that is

XX not present in a native form of the forestry plant

XX Sequence 1293 BP; 359 A; 290 C; 313 G; 331 T; 0 U; 0 Other;

XX

XX

XX

Db	211	TCATATGGAATGGGGCGAGAGGAACCTGATTCGGGAAAACCTATGCGGACGACTTAAGGGGA	270
Oy	210	AGACCTTCTCAAGACTCTTGGACAAAGAGCTCTTAACGATTTGAGAGAGCTATCTTGT	269
Db	271	GGACCTCCTCAGAGGATTTGGACAGAGAACTTACCAATGATTTCCAGAGCGTGGTCT	330
Oy	270	GTGACCTTGGAAACCCGGTAGCCGTGATGCTTTATTTGGCTATGAGAGCTACAAAAGATG	329
Db	331	TTGGTCACCTTGATTCGGGCTGAACGTGATGCGTACTTTGGCGAATGAAAGGACGAAAAGATG	390
Oy	330	GACTTCAGAGCAACCAAGTTCTTATGGAAGTTGCTTGCACAMGACATCAACGAGCTCT	389
Db	391	GACTTCAGAGCAACCAAGTTCTTATGGAAGTTGCTTGCACAMGACATCAACGAGCTCT	450
Oy	390	TCACGCTAGGCAAGGCTTACCAATGCTCGGCTACAAAGAAAGCTCTTGAAGAGGAGCTTGCTCA	449
Db	451	TATGGCAAGACAGACATATCATGCCCGATACAAAGAAAGCTATGAAAGAGAGTGCTCA	510
Oy	450	CCACACTACCGGTGACTTCAGAAAGCTTTGGTTTCTCTGTTTACCTCATACAGTACGA	509
Db	511	CCACACACTGGAAGATTTTCGTAAGTTGGTGTACTCTTGGGAGCTCTACCGTAATGA	570
Oy	510	AGGAGATGAATGGAACATGACATTGGCTTAAGCAAGAGCTAAGCTGTGCATGAGAAAT	569
Db	571	TGGAAATAGGGAATATGACTTTGGCAAAAGCAGAGGCTAAGATATCCACAGAAAT	630
Oy	570	CAAGGACAGGACCTACAAATGATGAGAGATGTTATTTGAATCTTTGTCACAAAGAGCAAGC	629
Db	631	CTCAGAGAAAGGCTTATGGCCATGAGAGATCTCTAATGAATTTGGCTACTAGAGGAAAGC	690
Oy	630	TCAGATCAATGCTACTTTTAAACCGTTACCAAGATGATCATGGCGAGAAATTTCTCAAG	689
Db	691	ACAGGTCAATGCTACGCTGAATCACTACAAAATGATGTTGGAAATGATATCAACAAAGA	750
Oy	690	TCTTGAGGAAGAGATGATGATGACAAAGTTCTTGCACTTTGAGGTCAACCATTCAGTG	749
Db	751	TTT---GAAAACCTGATCCAAAAGACGCGTTCCTTACTATACGAGGCTACGTAAGTG	807
Oy	750	CTTGAACAGACCAAGGCTTACTTTGTGATGATCTTTCGTTCAGAAATCAACAAAACCTGG	809
Db	808	CCTGACTCGCCTTGAAAGATTTTGAAGAGTTCTTGTCTGACCATCAATTAACGAGG	867
Oy	810	AACTGATGAAGGACACTCACTAGAATTTGTGSCCAAGAGCTGAGATTTGAATTGAAGGT	869
Db	868	AACGATGAAGGGGCTCTGACCAAGATGTTGCTTACAGGGCCGAGGTGACATGAAGTT	927
Oy	870	CATTGAGAGAGATACCAAGCCGAGGAACAGACATTCCTTTGAGAAAGCTATTAACAAAGA	929
Db	928	TATTAAGTAGAGATACCAAGAGGAGAAATGACATCCTCTCGATCGTGCATTTGCAAGA	987
Oy	930	CACTCGTGGAATTAACGAAGATGCTGCTGCACTTCCGTGAAATGATGCTTAATC	989
Db	988	CACTACTGGAATATGAAGAAATGCTTCTGGCATTAATGACCACGTCGAGGCTTGATTT	1047
Oy	990	AATCAATCTTC 1000	
Db	1048	TACAAAGTACTC 1058	
RESULT 13			
ADB94943			
ID	ADB94943	standard; cDNA; 1293 BP.	
AC	ADB94943;		
XX	04-DEC-2003	(first entry)	
DE	Programmed cell death pathway protein annexin cDNA #1.		
KM	programmed cell death; plant development; plant cell cycle; AT12; DAD1;		
KM	Dnase; 11e; 14d1; nucellin-like aspartic protease; annexin; prohibitin;		
KM	fen-1-like protein; rsc2; retinoblastoma-related protein; SIN3; TFIID;		

KW	TEGT; xylogenic Rnaase; pur-alpha; cyteine proteasee; RPPs-like protein;
KW	gp 91 NADPH oxidase subunit; NPR-1like protein; BAG-1;
KW	defender against cell death; lethal leaf spot; lesion stimulating death;
KW	seven in absentia; transcription initiation factor;
KM	testis enhanced gene transcript; gene; ss.
XX	
OS	Pinus radiata.
PN	US2003082724-A1.
PD	01-MAY-2003.
PX	
PF	14-AUG-2002; 2002US-00219220.
PR	04-JUN-1999; 99US-00325932.
XX	
PA	(GENE-) GENESIS RES & DEV CORP LTD.
PI	Plum B, Iasham A;
DR	WPI; 2003-786916/74.
PT	New isolated polynucleotide useful for modulating programmed cell death,
PT	altering the development cycle of plant cells, and subsequently modifying
PS	plant development.
XX	
PS	Claim 1; Page 28-29; 214pp; English.
CC	
CC	The invention describes an isolated polynucleotide (I) comprising a
CC	nucleotide sequence that is one of 145 fully defined sequences of 221-
CC	3415 base pairs (bp), given in the specification, and/or its complements,
CC	reverse complements, reverse sequences, or sequences having 75, 90 or 95
CC	% sequence identity to or that hybridize under stringent hybridisation
CC	conditions to one of the 145 sequences. The methods and compositions of
CC	the present invention to do with (I) are useful for modulating programmed
CC	cell death and thereby altering the development cycle of plant cells, and
CC	altering plant development. This sequence encodes a protein associated
CC	with the programmed cell death pathway.
SQ	
SQ	Sequence 1293 BP; 359 A; 290 C; 313 G; 331 T; 0 U; 0 Other;
	Query Match 38.1%; Score 468.6; DB 10; Length 1293;
	Best Local Similarity 68.4%; Pred.No. 2.5e-109;
	Matches 664; Conservative 0; Mismatches 304; Indels 3; Gaps 1;
QY	
DB	30 AAAAATGGCACTCTTAAGTTCGATTCTGATTCGTCTCGTCTCTTGATGATGCTGAGCA 89
QY	91 AACATGTCGACTCTCACCGGCCGACGCACCTGCCCTCTGAGCCGATGACTGCGAGCA 150
QY	90 ATTGAAGAACCGCTTTTGAAGATGGGGTAGCAACAGAGAATTGATCATATCAATCTTGGC 149
DB	151 GCTCCGACAGACCTTCGCGAGATGGGAAACAAATGABAAGCTATCATATCATATTGGG 210
QY	150 TCACAGAAAGTCTGTAACAAGAGAAAGTCATCAGGCAAGCATACCAGAAACTTACGGCGA 209
DB	211 TCATAGGAATGCGCGCAGAGAAAGTCGATTCGGCAAAACCTATGCGAGACTTACGGCGA 270
QY	210 AGACCTTCTCAAGACTCTTGACAAGAGAGCTCTTAAAGATTTGAGAGAGCTATCTTGT 269
DB	271 GGACCTCTCCAAGGCAATTTGACAAAGAACTTAAACAAATGATTTCCAGAGGCGCTGGTGCCT 330
QY	270 GTGACCTCTTGAACCCGGTGGCGGTGATGCTTAAATTTGGCTAATGAACTCAAAAAAGATG 329
DB	331 TTGGTCACTTGATCGGCTGAAACGTGATGGGTACTTGGCGAAATGABACGACGAAAAAGTG 390
QY	330 GACTTCAAGCAACCAAGTTCTTATGAAAGTTGCTTGGCACAAGGACATCAAAGCAAGTGTCT 389
DB	391 GACTTCAAGCAACCAAGTTCTTATGAAAGTTGCTTGGCACAAGGACATCAAAGCAAGTGTCT 450
QY	390 TCACGCTTAGGCAAGCTTACCATGTGCTGCTTCAAGAAAGTCTCTTGAAGAGAGCGTTGCTCA 449
DB	451 TATGGCAAGACCAAGCATATCATGCCCCCATCAAGAAATGCTCAATGGAAGAGAGCGTGGCTCA 510



PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 26-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147320P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.

PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 23-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161929P.  
PR 28-OCT-1999; 99US-0161933P.  
PR 29-OCT-1999; 99US-0162142P.

Query Match 37.2%; Score 457.2; DB 3; Length 1160;  
Best Local Similarity 67.3%; Pred. No. 2e-106;  
Matches 645; Conservative 0; Mismatches 313; Indels 0; Gaps 0;

QY 30 AAAATGGGACCTTAAGGTTTCTGATTCGTGTCGGCTTCCTGATGATGCTGAGCA 89  
DB 92 AACATGGGCTCTCTCAAACTCCCAAGCAATGTTCTTCCGAAAGATGACCCGAGCA 151  
QY 90 ATTGAAACCGCTTTTGAAGATGGGTCAGAAAGAGCTTATCATATCAATCTTGGC 149  
DB 152 ACTCCAAAGGCTTTTCAGAGTGGGTACCAAGAGAGCTATCATATCAATCTTAC 211  
QY 150 TCACAGAGTGTGAAACAGAGAAATCATCAGGCAAGCTATCCAGAACTTACGGCCA 209  
DB 212 TCACAGAGAGCGACCAACGCGCTTGATCCGACGTTTAAAGCTTACCAATGA 271  
QY 210 AGACCTTTCACAGACTTTCGACCAAGAGGCTCTCTAAGATTTGAGAGAGCTATCTGTT 269  
DB 272 GGAATCTTTCACAGACTTTCGACCAAGAGGCTTTCACGACTTTGAGAGGCTGATGTT 331  
QY 270 GTGACTCTTGAACCCGATGAGCGTATGCTTATTTGGCTAATGAAGCTTACAAAAGATG 329  
DB 332 GTGACTCTTGAATCCACAGAGAGATGCTTATTTGGCTAAGATTCACCAAGATGTT 391  
QY 330 GACTTCAAGCAACCAAGTCTTATGGAAGTTGCTTGACAGAGCATCAACGAGCTGCT 389  
DB 392 CACCAAGAACAAATGGGTTCTTGTTGAAATCGCTTGACAAAGGCTTGCTTGAGCTTAT 451  
QY 390 TCACGCTAGGCAAGCTTACATGCTGCTGCTCAAGAGTCTTGAAGAGAGCTTGTCTCA 449  
DB 452 CAAGTCAAGCAAGCTTACAGCTGATACAGAAATCAATGAGAGAGATGTCGCGCA 511  
QY 450 CCACACTACCGTGAAGCTTGAAGAGCTTTGTTGTTCTTGTATCCTCATACAGATACGA 509  
DB 512 ACACACATCTGGTGAAGCTTGTAGCTTGTGCTCTCTTGAGACATTTACAGATATGA 571  
QY 510 AGGAGTGAAGTGAACATGACATTTGGCTTGAAGAGCTTAAAGCTGCTCAATGAGAAAT 569  
DB 572 AGGAGTGAAGTGAACATGATGCTTGCAAGATCTGAAAGCTTAAGATCTTACAGAGAGGT 631  
QY 570 CAAGGACAGCACTAATGATGAGGATGTTATTTGAATCTTGCAAGAGAGCAAGAC 629  
DB 632 CTCAGAGAAATCTTAAGTGAAGATGATCTTCAAGATCTTGAACAAAGAGAGCAAGC 691  
QY 630 TCAGATCAATGCTACTTTTAACTGTTACCAAGATGATCAAGAGAGAAATTCAGAG 689  
DB 692 ACAGCTCGGTGCAACCTACCACTACCAAGAGATGAAAGCGCATTACCAAGAA 751



Oy	690	TCTTTGAGGAAGAGATGATGACAAAGTTCCTTGCACTTTTGAGGTAAACCAATTCAGTG	7439
Db	752	CTTGAAGAAAGAGTCCGACGACATGACTCACTGAAGAACTCACTAAAGCTGTAAATCACATG	8111
Oy	750	CTTGACAAAGACCAAGCTTTACTTTGTCGATTTCTTCGTTCAGCAATCAACAAACTGG	8099
Db	812	TTTGACATACCTCGAGAGCAATTTTGAAAGGTTCTTGCTCATATCAATCAACAAATAGGG	8711
Oy	810	AATCGATGAAGAGGCACTCACTTGAATTTGTACCCAAAGAGTGTGATTTGACTTGAAGGT	8688
Db	872	AAACGAGATGGGGACTAACCCGATCGTCACTACAGAACTGAAGTTGACATGAAGACG	9311
Oy	870	CATTGAGAGAGATTAACAGCGACAGAAACAGACTTCCTTTGGAGAAAGCTATTACCAAGA	9299
Db	932	CATTAAGAGAGATATACAGGAAGAAACAGACTTCCTTTGACCGCTGTAATCGCCAAAGA	9911
Oy	930	CATCTGTGAGATTACGAGAGAGTCTGTGCACCTTCGGTGAAGATGATGCTTAA	987
Db	992	CACCTTCGTGTGCTAATGAGACATGCTTTGCTCTTCTTCGGACATGCGATGCTTGA	1049

RESULT 15  
AAC49728  
ID AAC49728 standard; DNA; 1156 BP.

AC AAC49728;

DT 18-OCT-2000 (first entry)

DE Arabidopsis thaliana DNA fragment SEQ ID NO: 62219.

KM Hybridisation assay; genetic mapping; gene expression control;  
KM protein identification; signal transduction pathway; metabolic pathway;  
KM promoter; termination sequence; ss.

OS *Arabidopsis thaliana*.

PN EP1033405-A2

PD 06-SEP-2000.

PF 25-FEB-2000, 2000EP-00301439.

PR 25-FEB-1999/ 99US-0121825P.

PR 09-MAR-1999) 99US-0123548P.

PR 25-MAR-1999 99US-0126264P.

PR 01-APR-1999 99US-0127462P.

PR 08-APR-1999 99US-0128714P.

PR 19-APR-1999 99US-0130077P.

PR 23-APR-1999; 99US-0130510P.

PR 28-APR-1999 99US-0131449P.

PR 30-APR-1999; 99US-0132407P.

PR 05-MAY-1999, 99US-0132485P.

PR 06-MAY-1999, 99US-0132487P.

PR 11-MAY-1999; 99US-0134256P.

PR 14-MAY-1999; 99US-0134219P.

PR 14-MAY-1999; 99US-0134370P.

PR 19-MAY-1999; 99US-0134941P.

PR 20-MAY-1999; 99US-0135124P.

PR	21-MAY-1999	99US-0135533Z
PR	24-MAY-1999	99US-0135629Z
PR	25-MAY-1999	99US-0136021P
PR	27-MAY-1999	99US-0136332P
PR	28-MAY-1999	99US-0136782P
PR	01-JUN-1999	99US-0137222P
PR	03-JUN-1999	99US-0137528P
PR	04-JUN-1999	99US-0137502P
PR	07-JUN-1999	99US-0137724P
PR	08-JUN-1999	99US-0138004P
PR	10-JUN-1999	99US-0138504Z
PR	14-JUN-1999	99US-0138847P
PR	14-JUN-1999	99US-0139119P
PR	16-JUN-1999	99US-0139452P
PR	16-JUN-1999	99US-0139453P
PR	17-JUN-1999	99US-0139458P
PR	18-JUN-1999	99US-0139459P
PR	18-JUN-1999	99US-0139462P
PR	18-JUN-1999	99US-0139463P
PR	18-JUN-1999	99US-0139463P
PR	18-JUN-1999	99US-0139763P
PR	21-JUN-1999	99US-0139817P
PR	22-JUN-1999	99US-0139899P
PR	23-JUN-1999	99US-0140353P
PR	23-JUN-1999	99US-0140354P
PR	24-JUN-1999	99US-0140695P
PR	24-JUN-1999	99US-0140823P
PR	29-JUN-1999	99US-0140991P
PR	30-JUN-1999	99US-0141287P
PR	01-JUL-1999	99US-0141842P
PR	01-JUL-1999	99US-0142154P
PR	02-JUL-1999	99US-0142055P
PR	06-JUL-1999	99US-0142330P
PR	08-JUL-1999	99US-0142803P
PR	09-JUL-1999	99US-0142920P
PR	12-JUL-1999	99US-0142977P
PR	13-JUL-1999	99US-0143524P
PR	14-JUL-1999	99US-0143624P
PR	15-JUL-1999	99US-0144005P
PR	16-JUL-1999	99US-0144085P
PR	16-JUL-1999	99US-0144086P
PR	19-JUL-1999	99US-0144335P
PR	19-JUL-1999	99US-0144331P
PR	19-JUL-1999	99US-0144332P
PR	19-JUL-1999	99US-0144333P
PR	19-JUL-1999	99US-0144334P
PR	19-JUL-1999	99US-0144335P
PR	22-JUL-1999	99US-0145086P
PR	22-JUL-1999	99US-0145087P
PR	22-JUL-1999	99US-0145089P
PR	22-JUL-1999	99US-0145192P
PR	23-JUL-1999	99US-0145145P
PR	23-JUL-1999	99US-0145218P
PR	26-JUL-1999	99US-0145224P
PR	26-JUL-1999	99US-0145276P
PR	27-JUL-1999	99US-0145913P
PR	27-JUL-1999	99US-0145918P
PR	27-JUL-1999	99US-0145919P
PR	28-JUL-1999	99US-0146386P
PR	28-JUL-1999	99US-0146386P



Tue Aug 23 09:03:46 2005

us-10-690-564-1.rng

Page 20

Search completed: August 22, 2005, 07:26:19  
Job time : 813 secs

---

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 22, 2005, 04:40:07 ; Search time 4649 Seconds

(without alignments)  
10070.779 Million cell updates/sec

Title: US-10-690-564-1

Perfect score: 1230  
1 cccgcgcgcgcgaacactaa.....aaaaaaaaaaaaaaaaaaaa 1230

Sequence: 1 cccgcgcgcgcgaacactaa.....aaaaaaaaaaaaaaaaaaaa 1230

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Database :

EST:\*  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_est3:\*  
4: gb\_est4:\*  
5: gb\_est5:\*  
6: gb\_est6:\*  
7: gb\_est7:\*  
8: gb\_est8:\*  
9: gb\_est9:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1108.6	90.1	1155	3	CNSOAB51
2	1081.4	87.9	1095	3	CNSOABR2
3	1058	86.0	1137	3	CNSOABTU
4	1030	83.7	1063	5	EX838972
5	947.6	77.0	966	3	CNSOABV3
6	924.2	75.1	1034	3	CNSOABY3M
7	892.2	72.5	910	7	CF652924
8	836.2	68.0	865	3	CNSOABCI1
9	831.8	67.6	843	7	CF652048
10	786.2	63.9	814	6	CA782049
11	743.6	60.5	782	5	BU635010
12	734.4	59.7	775	5	BU636350
13	703	57.2	703	1	AV784024
14	700	56.9	702	1	AV785121
15	667.8	54.3	679	5	EX836971
16	652.4	53.0	670	3	CNSOABT2
17	649.4	52.8	672	3	BP561140
18	641.4	52.1	660	6	CB256182
19	630.4	51.3	676	2	BE037587
20	627.2	51.0	661	5	BP560951
21	608.4	49.5	665	5	BP560725
22	601.4	48.9	627	1	AV782813
23	594	48.3	594	1	AV559022
24	590.4	48.0	592	1	AV538505

C	25	568.4	46.2	583	1	AI998553	AI998553	701546141
C	26	543.4	44.2	578	6	CB256316	CB256316	83-E01273
C	27	540.6	44.0	576	7	CF652799	CF652799	77-L02052
C	28	540.2	43.9	569	6	CB259325	CB259325	59-E9624-
C	29	538.6	43.8	720	6	CD815439	CD815439	BN15-.026F
C	30	535	43.5	535	9	CNSO058J	AL088081	Arabidops
C	31	512.2	41.6	667	6	CD816318	CD816318	BN15-.029D
C	32	505	41.1	676	6	CD818668	CD818668	BN20-.046E
C	33	504.2	41.0	673	6	CD829166	CD829166	BN40-.041E
C	34	502.4	40.8	518	6	CB257482	CB257482	63-E01179
C	35	500.4	40.7	767	6	CD814071	CD814071	BN15-.021O
C	36	500	40.7	500	1	AV550330	AV550330	AV550330
C	37	499.4	40.6	515	6	CB257484	CB257484	63-E01180
C	38	489	39.7	489	1	AV544263	AV544263	AV544263
C	39	488.8	39.7	496	1	CB256030	CB256030	07-E01166
C	40	485.8	39.5	929	6	CB350421	CB350421	P1F02 Cot
C	41	475.2	38.6	640	6	CD820443	CD820443	BN20-.052J
C	42	469	38.1	889	7	COL121757	COL121757	GR-EB03B
C	43	463	37.6	463	5	BP572264	BP572264	BP572264
C	44	463	37.6	464	5	BP581497	BP581497	BP581497
C	45	462.2	37.6	468	5	BP570388	BP570388	BP570388

#### ALIGNMENTS

RESULT 1  
CNSOAB51  
LOCUS  
DEFINITION  
CNSOAB51  
Arabidopsis thaliana Full-length cDNA Complete sequence from clone GS15TSL6Z805 of Silique of strain col-0 of Arabidopsis thaliana (thale cress).

ACCESSION  
BX818120  
VERSION  
BX818120.1  
KEYWORDS  
HTC; GS1T cDNA.  
SOURCE  
Arabidopsis thaliana  
ORGANISM  
Arabidopsis thaliana

REFERENCE  
AUTHORS  
Castelli, V., Aury, J.M., Jallion, O., Wincker, P., Clepet, C., Menard, M., Cruaud, C., Quetier, F., Scarpelli, C., Schacher, V., Temple, G., Caboche, M., Weissenbach, J., and Salanoubat, M.,  
Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences:  
A Combined Approach to Evaluate and Improve Arabidopsis Genome Annotation

#### TITLE

JOURNAL  
REFERENCE  
Unpublished  
2 (bases 1 to 1155)

AUTHORS  
Direct Submission  
Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :  
BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr  
- Web : www.genoscope.cns.fr)

#### COMMENT

The sequences are based on single pass reads.  
Life Technologies (a division of Invitrogen) members carried out full-length libraries construction : Temple G.  
Genoscope members carried out sequencing and annotation : Castelli V., Aury J.M., Jallion O., Wincker P., Menard M., Cruaud C., Schacher V., Weissenbach J., Salanoubat M.  
URG INRA : Clepet C., Caboche M.  
Annotation is based on the June 2003 version of the Arabidopsis genome released by MIPS (Munich Information center for Protein Sequences) . 5 prime and 3 prime are assembled with Phrap.  
http://www.genoscope.cns.fr/externe/sequences/Banque\_Projet\_EF/Full length  
http://www.genoscope.cns.fr/cgi-bin/ggb/ggbsource=Arabidopsis.

#### FEATURES

source  
1..1155  
Location/Qualifiers  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/strain="Col-0"  
/db\_xref="taxon:3702"

/clone="GSLTSL1622B06"  
/issue type="Silique"  
/plasmid="pCMVSPORT\_6"  
1..1155  
/gene="At1g35720"

## ORIGIN

Query Match 90.1%; Score 1108.6; DB 3; Length 1155;  
Best Local Similarity 98.6%; Pred. No. 1.1e-278;  
Matches 1139; Conservative 0; Mismatches 14; Indels 2; Gaps 2;

```
QY 21 AAGTAGAAGAAAAATGCGCACTCTTAAGGTTTCTGATTCCTGCTCTTCTGATGA 80
Db 1 AAGTAGAAGAAAAATGCGCACTCTTAAGGTTTCTGATTCCTGCTCTTCTGATGA 60

QY 81 TGTGTAGCAATTGAGAACCGCTTTTGAAGATGGGGTACGAAAGAGACTTGATCATTC 140
Db 61 TGTGTAGCAATTGAGAACCGCTTTTGAAGATGGGGTACGAAAGAGACTTGATCATTC 120

QY 141 AATCTTGGCTCAGAGAGTGTCTGAACAGAGAGAAAGTCATCAGGCAAGCATACCAAGAAAC 200
Db 121 AATCTTGGCTCAGAGAGTGTCTGAACAGAGAGAAAGTCATCAGGCAAGCATACCAAGAAAC 180

QY 201 CTACGCGAAGACCTTCTCAAGACTCTTGAACAAGAGCTCTTAACGATTTGAGAGAC 260
Db 181 CTACGCGAAGACCTTCTCAAGACTCTTGAACAAGAGCTCTTAACGATTTGAGAGAAC 240

QY 261 TATCTTGTGTGACCTTGAACCCGGTGAAGCGGTGATGCTTTATGGCTTAATGAACCTAC 320
Db 241 TATCTTGTGTGACCTTGAACCCGGTGAAGCGGTGATGCTTTATGGCTTAATGAACCTAC 300

QY 321 AAAAAGATGACTTCAAGCAACCAAGTCTTATGAAAGTTGTGCAACAAGACATCAAC 380
Db 301 AAAAAGATGACTTCAAGCAACCAAGTCTTATGAAAGTTGTGCAACAAGACATCAAC 360

QY 381 GCAGCTGCTTCAAGCTTGAAGCACTTACATGCTCTGCTACAGAAAGTCTTGAAGAGA 440
Db 361 GCAGCTGCTTCAAGCTTGAAGCACTTACATGCTCTGCTACAGAAAGTCTTGAAGAGA 420

QY 441 CGTTGTGACCAACATACCGGTGATGCTTGAAGGCTTTGCTTGTATACCTCATA 500
Db 421 CGTTGTGACCAACATACCGGTGATGCTTGAAGGCTTTGCTTGTATACCTCATA 480

QY 501 CAGGTACGAAGAGATGAAGTGAACATGATGCTTAAGCAAGAAGTAAAGCTGTGCA 560
Db 481 CAGGTACGAAGAGATGAAGTGAACATGATGCTTAAGCAAGAAGTAAAGCTGTGCA 540

QY 561 TGAAGAAATCAGAGACAGACATTAAGATGATGATGATGATGATGATGATGATGATG 620
Db 541 TGAAGAAATCAGAGACAGACATTAAGATGATGATGATGATGATGATGATGATGATG 600

QY 621 AAGCAAGCTCAGATCAATGCTCTTTTAAACGTTTCAAGATGATGATGATGATGATG 680
Db 601 AAGCAAGCTCAGATCAATGCTCTTTTAAACGTTTCAAGATGATGATGATGATGATG 660

QY 681 TCTCAAGAGCTTGAAGAGAGATGATGATGATGATGATGATGATGATGATGATGATG 740
Db 661 TCTCAAGAGCTTGAAGAGAGATGATGATGATGATGATGATGATGATGATGATGATG 720

QY 741 CATTCAGTGTGATCAAGACAGAGCTTTTCTTGTGATGTTCTTGTGATGATGATG 800
Db 721 CATTCAGTGTGATCAAGACAGAGCTTTTCTTGTGATGTTCTTGTGATGATGATG 780

QY 801 CAAAGCTGGAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 860
Db 781 CAAAGCTGGAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840

QY 861 CTGGAAGGT-CATTGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATG 919
Db 841 CTGGAAGGTCCATTGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATG 900

QY 920 TTACCAAGAGACTCGTGAAGATTAAGAGAAAGTCTCGTGAAGCTTTCTGATGAAGATG 979
Db 900 TTACCAAGAGACTCGTGAAGATTAAGAGAAAGTCTCGTGAAGCTTTCTGATGAAGATG 979
```

```
Db 901 TTACCAAAAAAACTGCGAGATTAAGAGAAAGTCTCTCGACACTTCGCTGAAGATG 960
QY 980 AAGCTTAATCATCATCTCCACAGAGAAACATAGAGCTCTACAGCTTCTGATATCT 1039
Db 961 AAGCTTAATCATCATCTCCACAGAGAAACATAGAGCTCTACAGCTTCTGATATCT 1020

QY 1040 CTTATCTCCCTCTCTCTCTCTTGTGATGATGATGATGATGATGATGATGATGATG 1099
Db 1021 CTTATCTCCCTCTCTCTCTCTTGTGATGATGATGATGATGATGATGATGATGATG 1080

QY 1100 ACCCTGTTGTTTCTGTTGTTGTTGATGATGATGATGATGATGATGATGATGATG 1158
Db 1081 ACCCTGTTGTTTCTGTTGTTGTTGATGATGATGATGATGATGATGATGATGATGATG 1140

QY 1159 AGAGAACCAAGGTGG 1173
Db 1141 AGAGAACCAAGGTGG 1155
```

RESULT 2  
CNS0ABR2  
LOCUS  
DEFINITION  
CNS0ABR2 1095 bp mRNA linear HTC 06-PEB-2004  
Arabidopsis thaliana Full-length cDNA Complete sequence from clone  
GSLTPEG56ZG07 of Hormone Treated Callus of strain col-0 of  
Arabidopsis thaliana (thale cress).

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
BX816579.1 GI:42471684  
HTC; GSLT; cDNA.  
Arabidopsis thaliana (thale cress)  
Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE  
AUTHORS  
TITLE  
1 (bases 1 to 1095)  
Castelli,V., Aury,J.M., Jaillon,O., Wincker,P., Clepet,C.,  
Menard,M., Cruaud,C., Queller,F., Scarpelli,C., Schachter,V.,  
Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.  
Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences:  
A Combined Approach to Evaluate and Improve Arabidopsis Genome  
Annotation

JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
Unpublished  
2 (bases 1 to 1095)  
Genoscope.  
Direct Submission  
Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :  
BP 191 91006 Evry cedex - FRANCE (E-mail : segre@genoscope.cns.fr)

COMMENT  
- Web : www.genoscope.cns.fr)  
The sequences are based on single pass reads.  
Life Technologies (a division of Invitrogen) members carried out  
full-length libraries construction : Temple G.  
Genoscope members carried out sequencing and annotation : Castelli  
V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,  
Schachter V., Weissenbach J., Salanoubat M.  
URGV INRA : Clepet C., Caboche M.  
Annotation is based on the June 2003 version of the Arabidopsis  
genome released by MIPS (Munich Information center for Protein  
Sequences) . 5 prime and 3 prime are assembled with Phrap.  
http://www.genoscope.cns.fr/externe/sequences/Banque\_Projet\_EF/Full  
length  
http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.

FEATURES  
source  
1..1095  
location/Qualifiers  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/strain="Col-0"  
/db\_xref="taxon:3702"  
/clone="GSLTPEG56ZG07"  
/issue type="Hormone Treated Callus"  
/plasmid="pCMVSPORT\_6"  
1..1095  
/gene="At1g35720"

ORIGIN

Query Match 87.9%; Score 1081.4; DB 3; Length 1095;  
 Best Local Similarity 99.8%; Pred. No. 1.4e-271;  
 Matches 1093; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 12 AAACACTAAAAGTGAAGAAAATGGGACTCTTAAAGTTTCTGATTCGTCTCTGCTCC 71  
 DB 1 AAACACTAAAAGTGAAGAAAATGGGACTCTTAAAGTTTCTGATTCGTCTCTGCTCC 60

QY 72 TTTCGATGATGCTTACGATTTGAGAACCCGCTTTTGAAGATGGGGTACGAACGAGACTT 131  
 DB 61 TTTCGATGATGCTTACGATTTGAGAACCCGCTTTTGAAGATGGGGTACGAACGAGACTT 120

QY 132 GATCATATCAATCTTGGCTCAGAGAGTGTGAACAGAGAAAGTCAATCAGGCAAGATA 191  
 DB 121 GATCATATCAATCTTGGCTCAGAGAGTGTGAACAGAGAAAGTCAATCAGGCAAGATA 180

QY 192 CCAAGAAACCTACGGCGAAGACCTTCTCAAGACTCTTGACAGAGAGCTCTTACGATTT 251  
 DB 181 CCAAGAAACCTACGGCGAAGACCTTCTCAAGACTCTTGACAGAGAGCTCTTACGATTT 240

QY 252 CGAGAGAGCTATCTTGTGTGAGACTCTTGAAACCCGGTGAAGCTGATCTTATTTGGCTAA 311  
 DB 241 CGAGAGAGCTATCTTGTGTGAGACTCTTGAAACCCGGTGAAGCTGATCTTATTTGGCTAA 300

QY 312 TGAAGCTACAAAAGATGAGACTTCAAGCAACCAAGTTCTTATGGAAGTTGGTGCACAG 371  
 DB 301 TGAAGCTACAAAAGATGAGACTTCAAGCAACCAAGTTCTTATGGAAGTTGGTGCACAG 360

QY 372 GACATCAACGAGCTGTCTTCAACGCTAGCAAGCTTACATGCTGCTTACAGAGAGTCTCT 431  
 DB 361 GACATCAACGAGCTGTCTTCAACGCTAGCAAGCTTACATGCTGCTTACAGAGAGTCTCT 420

QY 432 TGAAGAGAGAGCTGTCTCAACCACTACCGGTGATCTTCAAGAAAGCTTTTGGTTCTCTGT 491  
 DB 421 TGAAGAGAGAGCTGTCTCAACCACTACCGGTGATCTTCAAGAAAGCTTTTGGTTCTCTGT 480

QY 492 TACCTCATACAGGTGACGAAGAGATGGAAGTGAACATGACATTTGGCTTAAAGCAAGATA 551  
 DB 481 TACCTCATACAGGTGACGAAGAGATGGAAGTGAACATGACATTTGGCTTAAAGCAAGATA 540

QY 552 GCTGTGTCATGAGAAAATCAAGAGCAAGCACTACATGATGAGAGTATTTTGAATCTT 611  
 DB 541 GCTGTGTCATGAGAAAATCAAGAGCAAGCACTACATGATGAGAGTATTTTGAATCTT 600

QY 612 GTCCACAAGAGCAAGAGCTGATCAATGCTACTTTTAAACCGTTTACCAAGATGATCAG 671  
 DB 601 GTCCACAAGAGCAAGAGCTGATCAATGCTACTTTTAAACCGTTTACCAAGATGATCAG 660

QY 672 CGAGAAATCTCAAGAGCTTGAAGAGAGATGATGAGCAAAAGTTCTTGTGACTTTT 731  
 DB 661 CGAGAAATCTCAAGAGCTTGAAGAGAGATGATGAGCAAAAGTTCTTGTGACTTTT 720

QY 732 GAGGTCAACCATTCAGTGTGACAAAGACAGAGCTTACTTTGTGATGTTCTTCTGCTTC 791  
 DB 721 GAGGTCAACCATTCAGTGTGACAAAGACAGAGCTTACTTTGTGATGTTCTTCTGCTTC 780

QY 792 AGCAATCAACAAAATCTGAACTGATGGAAGAGCACTCACTAAGATTTGTGACCAAGAGC 851  
 DB 781 AGCAATCAACAAAATCTGAACTGATGGAAGAGCACTCACTAAGATTTGTGACCAAGAGC 840

QY 852 TGAAGATGACTTGAAGGTCAATTGGAAGAGGATCAAGCGAGAAACGACTTCTTGTGA 911  
 DB 841 TGAAGATGACTTGAAGGTCAATTGGAAGAGGATCAAGCGAGAAACGACTTCTTGTGA 900

QY 912 GAAAGCTATTTACCAAGACACTCGTGAAGATTTACGAGAAATGCTCTGCACTTCTCG 971  
 DB 901 GAAAGCTATTTACCAAGACACTCGTGAAGATTTACGAGAAATGCTCTGCACTTCTCG 960

QY 972 TGAAGATGATGCTTATCAATCAATCTCTCAACAGAGAAACATTAAGTCTCTACAGTTCT 1031  
 DB 961 TGAAGATGATGCTTATCAATCAATCTCTCAACAGAGAAACATTAAGTCTCTACAGTTCT 1020

QY 1032 TGTATATCTTATCTCCCTCTCTCT-CTTTGATGAGTTTCAAAATCGTTTGAATTTGTT 1090

Db 1021 TGTATATCTTATCTCCCTATCTCTCTGCTTATGATGATTCAAAATCGTTTGAATTTGTT 1080

QY 1091 TCTACAAAACCTTG 1105

Db 1081 TCTACAAAACCTTG 1095

RESULT 3  
 CDSABTU 1137 bp mRNA linear HTC 06-FEB-2004  
 LOCUS Arabidopsis thaliana Full-length cDNA Complete sequence from clone  
 DEFINITION GSUTGH182F02 of Hormone Treated Callus of strain col-0 of  
 Arabidopsis thaliana (thale cress).  
 BX815936  
 ACCESSION BX815936.1 GI:42472184  
 VERSION HTC; GSUT cDNA.  
 KEYWORDS Arabidopsis thaliana (thale cress)  
 SOURCE Arabidopsis thaliana  
 ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
 1 (bases 1 to 1137)  
 Castella V., Aury J.M., Jallion O., Wincker P., Clepet C.,  
 Menard M., Chaud C., Quetier F., Scarpelli C., Schachter V.,  
 Temple G., Caboche M., Weisenbach J. and Salanoubat M.  
 Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences:  
 A Combined Approach to Evaluate and Improve Arabidopsis Genome  
 Annotation  
 Unpublished  
 2 (bases 1 to 1137)

JOURNAL Genoscope.  
 REFERENCE Direct Submission  
 AUTHORS Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :  
 TITLE BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr  
 JOURNAL - Web : www.genoscope.cns.fr)  
 COMMENT The sequences are based on single pass reads.  
 Life Technologies (a division of Invitrogen) members carried out  
 full-length libraries construction : Temple G.  
 Genoscope members carried out sequencing and annotation : Castella  
 V., Aury J.M., Jallion O., Wincker P., Menard M., Chaud C.,  
 Schachter V., Weisenbach J., Salanoubat M.  
 URGV INRA : Clepet C., Caboche M.  
 Annotation is based on the June 2003 version of the Arabidopsis  
 genome released by MIPS (Munich Information Center for Protein  
 Sequences) : 5 prime and 3 prime are assembled with phrap.  
 http://www.genoscope.cns.fr/externe/sequences/Banque\_Projet\_BP/Full  
 length  
 http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.  
 FEATURES  
 source Location/Qualifiers  
 1. 1137  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /strain="Col-0"  
 /db\_xref="taxon:3702"  
 /clone="GSUTGH182F02"  
 /issue\_type="Hormone Treated Callus"  
 /plasmid="pCMVSPORT\_6"  
 1. 1137  
 /gene="At1g35720"

ORIGIN  
 Query Match 86.0%; Score 1058; DB 3; Length 1137;  
 Best Local Similarity 97.8%; Pred. No. 1.9e-265;  
 Matches 1105; Conservative 0; Mismatches 20; Indels 5; Gaps 3;

QY 12 AAACACTAAAAGTGAAGAAAATGGS-CCGACTCTTAAGGT-TCGTATTCGTCTCCG 67  
 DB 2 AAACACTAAAAGTGAAGAAAATGAGAACTCTTAAGGAATTAAGAAATCTGTTCCG 61

QY 68 CTCCTCTGATGATGCTGAGCAATTGAGAACCGCTTTTGAAGATGGGTACGAAGAG 127  
 DB 62 CTCCTCTGATGATGCTGAGCAATTGAGAACCGCTTTTGAAGATGGGTACGAAGAG 121

QY	128	ACTGATCATATCAATCTTTGGCTCACAAAGTGTGAAACAGAGAAATCATCAAGGCAAG	187
QY	128	ACTGATCATATCAATCTTTGGCTCACAAAGTGTGAAACAGAGAAATCATCAAGGCAAG <td>187</td>	187
Db	122	ACTGATCATATCAATCTTTGGCTCACAAAGTGTGAAACAGAGAAATCATCAAGGCAAG	181
QY	188	CATACCAAGAAACCTACGGCGGAGACCTTCTCAAGACTCTTGACAGAGAGCTCTTAAACG	247
Db	182	CATACCAAGAAACCTACGGCGGAGACCTTCTCAAGACTCTTGACAGAGAGCTCTTAAACG	241
QY	248	ATTTCGAGAGAGCTATCTTTGTTGTGAGACTCTTGAAACCCGGGTGAGGCTGATGCTTTATTGG	307
Db	242	ATTTCGAGAGAGCTATCTTTGTTGTGAGACTCTTGAAACCCGGGTGAGGCTGATGCTTTATTGG	301
QY	308	CTAATGAAGCTACAAAAGATGGACTTCAGACCAACCAAGTCTTATGGAAGTGCTTGCA	367
Db	302	CTAATGAAGCTACAAAAGATGGACTTCAGACCAACCAAGTCTTATGGAAGTGCTTGCA	361
QY	368	CAAGGACATCAACCGAGCTGCTTCCAGTACGCAAGCTTACCATGCTCTGCTACAAAGAGT	427
Db	362	CAAGGACATCAACCGAGCTGCTTCCAGTACGCAAGCTTACCATGCTCTGCTACAAAGAGT	421
QY	428	CTCTTGAAAGAGAGCTTGGCTCACCAACTACCGGTGACTTCAGAAAAGCTTTTGGTTTCTC	487
Db	422	CTCTTGAAAGAGAGCTTGGCTCACCAACTACCGGTGACTTCAGAAAAGCTTTTGGTTTCTC	481
QY	488	TTGTTACTCTCATACAGGTACGAGAGAGATGGAATGATGACATTTGCTTAAGCAAGAG	547
Db	482	TTGTTACTCTCATACAGGTACGAGAGAGATGGAATGATGACATTTGCTTAAGCAAGAG	541
QY	548	CTAAGCTGTCCATGAGAAAATCAAGGACAGCACTACATATGATGAGATGTTATTGAA	607
Db	542	CTAAGCTGTCCATGAGAAAATCAAGGACAGCACTACATATGATGAGATGTTATTGAA	601
QY	608	TCTTGTCCACAAGAGGCAAGCTCAGATCAATGCTATCTTTTAAACCGTTAACCAAGATGATC	667
Db	602	TCTTGTCCACAAGAGGCAAGCTCAGATCAATGCTATCTTTTAAACCGTTAACCAAGATGATC	661
QY	668	ATGGCGAGAAATTTCTCAAGAGTCTTGAGAGAGAGATGATGATGACAAGTTCCTTGGAC	727
Db	662	ATGGCGAGAAATTTCTCAAGAGTCTTGAGAGAGAGATGATGATGACAAGTTCCTTGGAC	721
QY	728	TTTTGAGGTCAACCATTCAGTGTCTTGACAAGACCGAGCTTTTACTTTGTGATGATGTTCTTC	787
Db	722	TTTTGAGGTCAACCATTCAGTGTCTTGACAAGACCGAGCTTTTACTTTGTGATGATGTTCTTC	781
QY	788	GTTACGCAATCAACAAACCTGGAATGTGATGGAAGGACCTCACTGATTTGTGACCAACA	847
Db	782	GTTACGCAATCAACAAACCTGGAATGTGATGGAAGGACCTCACTGATTTGTGACCAACA	841
QY	848	GAGCTGAATTTGACTTTGAAGGTCAATTGAGAGAGATACACAGGCGAGAAACAGATTTCTT	907
Db	842	GAGCTGAATTTGACTTTGAAGGTCAATTGAGAGAGATACACAGGCGAGAAACAGATTTCTT	901
QY	908	TGAGAGAAAGCTATTACCAAGACACTGTGTGAGATTACGAGAAGATGCTGTGCGACTTC	967
Db	902	TGAGAGAAAGCTATTACCAAGACACTGTGTGAGATTACGAGAAGATGCTGTGCGACTTC	961
QY	968	TGCGTGAAGATGATGCTTATATCAATCAATCCTCCACAGAGAAACATAGCTGCTCTAAG	1022
Db	962	TGCGTGAAGATGATGCTTATATCAATCAATCCTCCACAGAGAAACATAGCTGCTCTAAG	1021
QY	1028	CTTCTGTTATCTCTTATCTCCCTCTCTCTCTCTTGTGATGAGTTCAAAATCGTTGATTTT	1087
Db	1022	CTTATGTTATCTCTTATCAACCTGTGTCTATCTATGATGAGAGTTCAAAATCGTTGATTTT	1081
QY	1088	GTTTCTTACA-AAAACCTTGTGTTGTTCTGTGTGTGTGTTTGAAGTTCCTAA	1136
Db	1082	TTTTTTATATATAAACCTTGTGTTGTTCTGTGTGTGTGTTTGAAGTTCCTAA	1131

RESULT 4

EX838972

1063 bp

mRNA

linear

EST 11-FEB-2004

	DEFINITION
ACCESSION	BX38972 Arabidopsis thaliana Flowers and buds Col-0 Arabidopsis thaliana CDNA clone GSLTFB50Z03 5PRIM, mRNA sequence.
VERSION	BX38972
KEYWORDS	BX38972.1 GI:42533055
SOURCE	EST.
ORGANISM	Arabidopsis thaliana (thale cress)
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliopsida; eudicotyledons; core eudicots; rosids; eurosid II; Brassicales; Brassicaceae; Arabidopses.
AUTHORS	Castelli,V., Aury,J.M., Jallion,O., Winkler,P., Clepet,C., Menard,M., Crnaud,C., Quetier,F., Scarpelli,C., Schachter,V., Temple,G., Caboche,M., Weissenbach,J. and Salanoubat M. Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences A Combined Approach to Evaluate and Improve Arabidopsis Genome Annotation
TITLE	Unpublished (2004)
JOURNAL	Contact: Genoscope
COMMENT	Genosco - Centre National de Sequencage 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr The sequences are based on single pass reads. Life Technologies (a division of Invitrogen) members carried out full-length libraries construction : Temple G. Genoscope members carried out sequencing and annotation : Castelli V., Aury J.M., Jallion O., Winkler P., Menard M., Crnaud C., Schachter V., Weissenbach J., Salanoubat M. URUG INRA : Clepet C., Caboche M. Annotation is based on the June 2003 version of the Arabidopsis genome released by MIPS (Munich information center for Protein Sequences). <a href="http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/EST">http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/EST</a> <a href="http://www.genoscope.cns.fr/cgi-bin/g9b/g9bresource-Arabidopses">http://www.genoscope.cns.fr/cgi-bin/g9b/g9bresource-Arabidopses</a> .
FEATURES	source
	1..1063
	/organism="Arabidopsis thaliana"
	/mol_type="mRNA"
	/ecotype="Col-0"
	/db_xref="taxon:3702"
	/clone="GSLTFB50Z03"
	/tissue_type="Flowers and buds"
	/clone_lib="Arabidopsis thaliana Flowers and buds Col-0"
ORIGIN	
Query Match	83.7%; Score 1030; DB 5; Length 1063;
Best Local Similarity	99.0%; Pred. No.3,9e-258;
Matches 1036; Conservative	0; Mismatches 10; Indels 0; Gaps 0
QY	108 AGATGGGGTTCGAAGCAGACTTATCATATCAATCTTGCGTCAAGAGGTGTGAACA 167
DB	1 AGATGGGGTTCGAAGCAGACTTATCATATCAATCTTGCGTCAAGAGGTGTGAACA 60
QY	168 GAGAAAGTCATCAGGCAGCAATCACGAAACTTAGCGCGAGAACCCTTTCAAGACTCT 227
DB	61 GAGAAAGTCATCAGGCAGCAATCACGAAACTTAGCGCGAGAACCCTTTCAAGACTCT 120
QY	228 TGACAGAGAGCTCTTAACGATTGTGAGAGAGCTATCTTTGTGTGACCTTTGAACCCGG 287
DB	121 TGACAGAGAGCTCTTAAAGATTGTGAGAGAGCTATCTTTGTGTGACCTTTGAACCCGG 180
QY	288 TGAGCGTATGCTTTATTTGGCTTAATGAAGCTCAAAAAAGATGATGACTTCAAGCAACCAAT 347
DB	161 TGAGCGTATGCTTTATTTGGCTTAATGAAGCTCAAAAAAGATGATGACTTCAAGCAACCAAT 240
QY	348 TCTTATGGAAGTTGCTTGCACAGAGCAATCAACGAGCTGCTTACGCTAGGCCAAGCTTA 407
DB	241 TCTTATGGAAGTTGCTTGCACAGAGCAATCAACGAGCTGCTTACGCTAGGCCAAGCTTA 300
QY	408 CCATGCTGGCTACAGAAAGTCTTTGGAAGAGACCTTGCTCCACACATACCGGTGACTT 467
DB	301 CCATGCTGGCTACAGAAAGTCTTTGGAAGAGACCTTGCTCCACACATACCGGTGACTT 360

QY 468 CAGAAAGCTTTGGTTCTCTGTTACTCTATACAGATGAGAGAGATGAGTGAACAT 527  
 DB 361 CAGAAAGCTTTGGTTCTCTGTTACTCTATACAGATGAGAGAGATGAGTGAACAT 420  
 QY 528 GACATTTGGCTTAAGCAAGAGAGTAACTGGTCCATGAGAAATCAAGACCAAGACTCA 587  
 DB 421 GACATTTGGCTTAAGCAAGAGAGTAACTGGTCCATGAGAAATCAAGACCAAGACTCA 480  
 QY 588 TGAATGAGATGTTATTAGAAATCTTGTCCACAAAGAGCAAACTGATCAATGCTACTT 647  
 DB 481 TGAATGAGATGTTATTAGAAATCTTGTCCACAAAGAGCAAACTGATCAATGCTACTT 540  
 QY 648 TAAACCTTACCAAGATGATCATGAGGAGAAATCTCAAGAGTCTTGAGAGAGATGA 707  
 DB 541 TAAACCTTACCAAGATGATCATGAGGAGAAATCTCAAGAGTCTTGAGAGAGATGA 600  
 QY 708 TGAATGACAGTTCTTGTGACCTTTGAGGTCAACCAATTCAGTGTGACCAAGACCA 767  
 DB 601 TGAATGACAGTTCTTGTGACCTTTGAGGTCAACCAATTCAGTGTGACCAAGACCA 660  
 QY 768 TTACTTTGTCATGTTCTTGTGTCAGCAATCAACAACTGGAATGAGAGACACT 827  
 DB 661 TTACTTTGTCATGTTCTTGTGTCAGCAATCAACAACTGGAATGAGAGACACT 720  
 QY 828 CACTAGAAATGTCACACAGAGCTGAGATTGATGAGAGGTCAATGAGAGAGTCCA 887  
 DB 721 CACTAGAAATGTCACACAGAGCTGAGATTGATGAGAGGTCAATGAGAGAGTCCA 780  
 QY 888 GCGGAGAGAAAGCAATTCCTTTGAGAAAGCTATTACCAAGACACTGAGAGATTGA 947  
 DB 781 GCGGAGAGAAAGCAATTCCTTTGAGAAAGCTATTACCAAGACACTGAGAGATTGA 840  
 QY 948 GAAATGTCGTGTCGACCTTCTCGGTGAAGATGATGCTTAATCAATCAATCTTCA 1007  
 DB 841 GAAATGTCGTGTCGACCTTCTCGGTGAAGATGATGCTTAATCAATCAATCTTCA 900  
 QY 1008 AAACATAAGCTGCTCTACAGCTTCGTATCTTATCTCCCTCTCTCTTGAATGA 1067  
 DB 901 AAACATAAGCTGCTCTACAGCTTCGTATCTTATCTCCCTCTCTCTTGAATGA 960  
 QY 1068 GTTTCAAATGCTTTGATTTGTTTCTACAAAAACCTGTTGTTGTTGTTGTTG 1127  
 DB 961 GTTTCAAATGCTTTGATTTGTTTCTAAAAAAACCTGTTGTTGTTGTTGTTG 1020  
 QY 1128 AGTTCCTTAATTAAGCAAGAGAGA 1153  
 DB 1021 AGTTCCTTAATTAAGCAAGAGAGA 1046

RESULT 5  
 CDSABV3 966 bp mRNA linear HTC 06-FEB-2004  
 LOCUS Arabidopsis thaliana Full-length cDNA Complete sequence from clone  
 DEFINITION GILTPGH18ZH06 of Hormone Treated Callus of strain col-0 of  
 Arabidopsis thaliana (thale cress).  
 BX815939  
 ACCESSION BX815939.1 GI:42472187  
 VERSION HTC; GSI1\_cDNA  
 SOURCE Arabidopsis thaliana (thale cress)  
 KEYWORDS Arabidopsis thaliana  
 ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis  
 1 (bases 1 to 966)  
 Castelli,V., Aury,J.M., Jallion,O., Wincker,P., Clepet,C.,  
 Menard,M., Cruaud,C., Quecier,F., Scarpelli,C., Schachter,V.,  
 Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.  
 TITLE Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences:  
 A Combined Approach to Evaluate and Improve Arabidopsis Genome  
 Annotation  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 966)  
 AUTHORS Genoscope.

TITLE Direct Submission  
 JOURNAL Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :  
 BP 191 91006 Evry cedex - FRANCE (E-mail : seqref@genoscope.cns.fr  
 - Web : www.genoscope.cns.fr)  
 COMMENT The sequences are based on single pass reads.  
 Life Technologies (a division of Invitrogen) members carried out  
 full-length libraries construction : Temple G.  
 Genoscope members carried out sequencing and annotation : Castelli  
 V., Aury J.M., Jallion O., Wincker P., Menard M., Cruaud C.,  
 Schachter V., Weissenbach J., Salanoubat M.  
 URG INRA : Clepet C., Caboche M.  
 Annotation is based on the June 2003 version of the Arabidopsis  
 genome released by MIPS (Munich Information center for Protein  
 Sequences). 5 prime and 3 prime are assembled with Phrap.  
 http://www.genoscope.cns.fr/externe/sequences/Banque\_Projet\_EF/Full  
 length  
 http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.  
 FEATURES  
 source Location/Qualifiers  
 1..966  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /strain="Col-0"  
 /db\_xref="taxon:3702"  
 /clone="GSLRPGH18ZH06"  
 /issue\_type="Hormone Treated Callus"  
 /plasmid="PCWVS-PORT\_6"  
 1..966  
 /gene="At1g35720"  
 ORIGIN  
 Query Match 77.0%; Score 947.6; DB 3; Length 966;  
 Best Local Similarity 99.5%; Pred. No. 1,4e-236;  
 Matches 961; Conservative 0; Mismatches 4; Indels 1; Gaps 1;  
 QY 195 CGAAACCTACGGGAGAAACCTTCTCAAGACTCTTGAACAGAGACTCTACATTTGCA 254  
 DB 1 CGAGACTTACGGGAGAAACCTTCTCAAGACTCTTGAACAGAGACTCTACATTTGCA 60  
 QY 255 GAGAGCTATCTTGTGTGACTCTTGAACCCGGTGAAGCTGATGCTTTATTTGGCTAATGA 314  
 DB 61 GAGAGCTATCTTGTGTGACTCTTGAACCCGGTGAAGCTGATGCTTTATTTGGCTAATGA 120  
 QY 315 AGCTACAAAGATGAGACTTCAAGCAACAAAGTTCTTATGAGAGTCTTGCAAGAGAC 374  
 DB 121 AGCTACAAAGATGAGACTTCAAGCAACAAAGTTCTTATGAGAGTCTTGCAAGAGAC 180  
 QY 375 ATCAACGACAGCTGCTTCAAGCTGAGCAAGCTTACATGCTGCTACCAAGAGTCTTGA 434  
 DB 181 ATCAACGACAGCTGCTTCAAGCTGAGCAAGCTTACATGCTGCTACCAAGAGTCTTGA 240  
 QY 435 AGAGAGCTTGTCTACACACTACCGGTGACTTCAAGAACTTTGGTTCTCTGTTAC 494  
 DB 241 AGAGAGCTTGTCTACACACTACCGGTGACTTCAAGAACTTTGGTTCTCTGTTAC 300  
 QY 495 CTGATACAGGTACGAGAGATGAATGAATGACATTTGGCTTACGACAGAGACTTAAGCT 554  
 DB 301 CTGATACAGGTACGAGAGATGAATGAATGACATTTGGCTTACGACAGAGACTTAAGCT 360  
 QY 555 GGTCCATGAGAAATCAAGAGCAAGCACTACATGATGATGAGATGATTAATGATCTTGTG 614  
 DB 361 GGTCCATGAGAAATCAAGAGCAAGCACTACATGATGATGAGATGATTAATGATCTTGTG 420  
 QY 615 CACAAGAGCAAGCTCAGATCAATGCTATTTTAAACGTTACCAAGATGATGAGCGCA 674  
 DB 421 CACAAGAGCAAGCTCAGATCAATGCTATTTTAAACGTTACCAAGATGATGAGCGCA 480  
 QY 675 GGAATTTCCAAAGAGCTTGAAGAGATGATGATGACAAAGTCTTGCACTTTTGAAG 734  
 DB 481 GGAATTTCCAAAGAGCTTGAAGAGATGATGATGACAAAGTCTTGCACTTTTGAAG 540  
 QY 735 GTCAACATTCAGTGTGATGACCAAGACCAAGAGCTTACTTTGTCATGATGTTCTGTCACG 794  
 DB 541 GTCAACATTCAGTGTGATGACCAAGACCAAGAGCTTACTTTGTCATGATGTTCTGTCACG 600



QY	795	AATCAACAAAAC	CTGATGATGAGAGCACTCACTGAATTGATGACCAAGAGCTGA	854
Db	601	AATCAACAAAAC	CTGATGATGAGAGCACTCACTGAATTGATGACCAAGAGCTGA	660
QY	855	GATTGACTTGAAGT	CATTGGAGAGAGTACCGAGCAAGCAAGCATTTCTTTGGAGAA	914
Db	661	GATTGACTTGAAGT	CATTGGAGAGAGTACCGAGCAAGCAAGCATTTCTTTGGAGAA	720
QY	915	AGCTATTACAAAG	CACCTCGTGGAGATTACGAAAGATGCTTCGTGCACTTTCGGTGA	974
Db	721	AGCTATTACAAAG	CACCTCGTGGAGATTACGAAAGATGCTTCGTGCACTTTCGGTGA	780
QY	975	AGATGATGCTTAA	TATCAATCAATCCCTCCACAGAGAAACATAAGCTGCTCAAGCTTCTGT	1034
Db	761	AGATGATGCTTAA	TATCAATCAATCCCTCCACAGAGAAACATAAGCTGCTCAAGCTTCTGT	840
QY	1035	TATCTCTTATCT	CCCTCTCTCTCTCTTGTGATGATGATTTCAAAATCGTTTGAATTTTGTCTA	1094
Db	841	TATCTCTTATCT	CCCTCTCTCTCTCTTGTGATGATGATTTCAAAATCGTTTGAATTTTGTCTA	900
QY	1095	CAAAA	ACCTTGTTGTTGTTCTGTGTGTGTTTGATTTGATTTCTTAATTAATGCAAAAGAGAA	1155
Db	901	CAAAAACACCT	TGTTGTTGTTCTGTGTGTGTTTGATTTGATTTCTTAATTAATGCAAAAGAGAA	960
QY	1154	GACAGA	1159	
Db	961	GAAAGA	966	
RESULT 6				
CNS09Y3M				
LOCUS		1034 bp	mRNA	linear
DEFINITION		Arabidopsis thaliana full-length cDNA	Complete sequence	from clone
		GSU1833ZB01	of Adult vegetative tissue	of strain col-0 of
		Arabidopsis thaliana	(thale cress).	
		BX841810		
ACCESSION		BX841810.1	GI:42454468	
VERSION		HTC; GSU1	cdna.	
KEYWORDS		Arabidopsis thaliana	(thale cress)	
ORGANISM		Arabidopsis thaliana		
		Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosid II; Brassicales; Brassicaceae; Arabidopsis.		
		1 (bases 1 to 1034)		
REFERENCE		Castelli,V., Aury,J.M., Jallion,O., Winkler,P., Clepet,C., Menard,M., Craud,C., Quetier,F., Scarpelli,C., Schachter,V., Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.		
AUTHORS		Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences: A Combined Approach to Evaluate and Improve Arabidopsis Genome Annotation		
TITLE		Unpublished		
JOURNAL		2 (bases 1 to 1034)		
REFERENCE		Genoscope.		
AUTHORS		Direct Submission		
TITLE		Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :		
JOURNAL		BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr		
		- Web : www.genoscope.cns.fr)		
COMMENT		The sequences are based on single pass reads.		
		Life technologies (a division of Invitrogen) members carried out		
		full-length libraries construction : Temple G.		
		Genoscope members carried out sequencing and annotation : Castelli		
		V., Aury J.M., Jallion O., Winkler P., Menard M., Craud C.,		
		Schachter V., Weissenbach J., Salanoubat M.		
		URCV INRA : Clepet C., Caboche M.		
		Annotation is based on the June 2003 version of the Arabidopsis		
		genome released by MIPS (Munich Information center for Protein		
		Sequences). 5 prime and 3 prime are assembled with phrap.		
		http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Fulllength		
		http://www.genoscope.cns.fr/cgi-bin/g9b/g9b?source=Arabidopsis.		
FEATURES		Location/Qualifiers		
SOURCE		1..1034		

		/organism="Arabidopsis thaliana"
		/mol_type="mRNA"
		/strain="Col-0"
		/db_xref="taxon:3702"
		/clone="GSLTIS23ZB01"
		/tissue_type="Adult vegetative tissue"
		/plasmid="PCWMSPORT_6"
		1..1034
		/gene="At1g35720"
ORIGIN		
Query Match	75.1%; Score 924.2; DB 3; Length 1034;	
Beet Local Similarity	98.0%; Pred. No. 1,9e-230;	
Matches	946; Conservative 0; Mismatches 18; Indels 1; Gaps 1	
OY	12 AAACCTAAAGTAGAAGAAAAATGGCACTCTTAAGGTTCGTATTCTGTCTCGCTCC	71
DB	1 AAACCTAAAGTAGAAGAAATGGCACTCTTAAGGTTCGTATTCTGTCTCGCTCC	60
OY	72 TTCTGATGATGCTGAACAATTGAGAACCGCTTTTGAGAGATGGGGTACGAACGAGACTT	131
DB	61 TTCTGATGATGCTGAACAATTGAGAACCGCTTTTGAGAGATGGGGTACGAACGAGACTT	120
OY	132 GATCATATCATTCTTGGCTCACAGAAGGCTGGAACAGAGGAAGCATACGSCAGACATA	191
DB	121 GATCATATCATTCTTGGCTCACAGAAGGCTGGAACAGAGGAAGCATACGSCAGACATA	180
OY	192 CCACGAAACCTACGCGCGAAGACCTTCTCAAGACTCTTGACAAAGAGCTCTCTAACGATTT	251
DB	181 CCACGAGACCTACGCGCGAAGACCTTCTCAAGACTCTTGACAAAGAGCTCTCTAACGATTT	240
OY	252 CGAGAGACTATCTTGTGTGGAGCTCTTGAAACCGGTGACGCTGATGCTTTATTGGCTTAA	311
DB	241 CGAGAGACTATCTTGTGTGGAGCTCTTGAAACCGGTGACGCTGATGCTTTATTGGCTTAA	300
OY	312 TGAAGCTACAAAANAATGGAAGCTTACAGAACCAAGTTCTTAATGSAAGTTCCTTGACAAAG	371
DB	301 TGAAGCTACAAAAGATGAGACTTCAAGAACCAAGTTCTTAATGSAAGTTCCTTGACAAAG	360
OY	372 GACATCAACGAGCGCTTTCACGCTAGGCAAGCTTAACCATGCTCGCTCAAGAAAGTCTCT	431
DB	361 GACATCAACGAGCGCTTTCACGCTAGGCAAGCTTCAACCATGCTCGCTCAAGAAAGTCTCT	420
OY	432 TGAAGAGACGTTGCTCAACAACCTACCGGTGACTTCAAAAAGCTTTTGTTTTCTCTGT	491
DB	421 TGAAGAGACGTTGCTCAACAACCTACCGGTGACTTCAAGAGCTTTTGTTTTCTCTGT	480
OY	492 TAACCTATACAGGTTCGAAGAGATGGAAGTGAACATGCATTGGCTTAAGCAAAGACTTAA	551
DB	481 TAACCTATACAGGTTCGAAGAGATGGAAGTGAACATGCATTGGCTTAAGCAAAGACTTAA	540
OY	552 GCTGTCATGAGAAATCAAGAACCAACATCAATGATGAGATGTTATTAGAACTTT	611
DB	541 GCTGTCATGAGAGATCAAGAGACACATCAATGATGAGATGTTATTAGAACTTT	600
OY	612 GTCCAAGAAAGCAAGCTCAGATCAATGCTTAAACCGTTTCCAAGATGATCATGG	671
DB	601 GTCCAAGAAAGCAAGCTCAGATCAATGCTTAAACCGTTTCCAAGATGATCATGG	660
OY	672 CGAGGAAATTTCTCAAGAGCTTTGAGGAAAGAGATGATGATGACAAAGTTCTTGCACTTTT	731
DB	661 CGAGGAAATTTCTCAAGAGCTTTGAGGAAAGAGATGATGATGACAAAGTTCTTGCACTTTT	720
OY	732 GAGGTCAACATTCAGGCTTGACAAGACCAAGGCTTAACTTTGTCAGATGTTCTTGCTTC	791
DB	721 GAGGTCAACATTCAGGCTTGACAAGACCAAGGCTTAACTTTGTCAGATGTTCTTGCTTC	780
OY	792 AGCAATCAACAAAATCTGGAAGTGAAGAGAGACTCACTGAATTTGTGACCACAAGAGC	851
DB	781 AGCAATCAACAAAATCTGGAAGTGAAGAGAGACTCACTGAATTTGTGACCACAAGAGC	840
OY	852 TGAAGTGAATTTGAAGTCAATGAGAGAGATCAAGGCGCAAGAAACAGCATTTCTTTGGA	911

Db 841 TGAGATTGACTTGAAAGCTGATTGAGAGGATACAGAGCAGAAACGACTTCCTTTGA 900

Qy 912 GAAAGCTATTACCAAGACACTCGTGGAGATTACGAGAGATGCTCGTGGACTTCGG 971

Db 901 GAGGCTATTACCAAGACACTCGTGGAGATTACGAGAGAG-GCTCGTGGCACTACTCGA 959

Qy 972 TGAAG 976

Db 960 TGAGG 964

RESULT 7  
CF652924 910 bp mRNA linear EST 06-NOV-2003  
LOCUS 85-L020135w-066-001-121-SP6P MP1Z-ADIS-066 Arabidopsis thaliana  
DEFINITION cDNA clone MP1Z20011211Q 5-PRIME, mRNA sequence.  
ACCESSION CF652924  
VERSION CF652924.1 GI:37429878  
KEYWORDS EST.  
SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana

REFERENCE  
AUTHORS Schmidt,K.U., Soerensen,T.R., Stracke,R., Torjek,O., Altmann,T.,  
Mitschke,I., Olda,T. and Weisshaar,B.  
TITLE Large-scale identification and analysis of genome-wide  
single-nucleotide polymorphisms for mapping in Arabidopsis thaliana  
Genome Res. 13 (6), 1250-1257 (2003)

JOURNAL MEDLINE 22683290  
PUBMED 12799357

COMMENT Contact: Weisshaar B  
ADIS DNA core facility at MP1Z  
Max-Planck-Institute for Plant Breeding Research  
Carl-von-Linne Weg 10, 50829 Koeln, Germany  
Fax: 00492215062851  
Email: weisshaar@mpiz-koeln.mpg.de  
Insert Length: 910 Std Error: 0.00  
Plate: 1 row: 1 column: 21  
Seq Primer: SP6P.  
Location/Qualifiers  
1. 910  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/ecotype="Ws-0"  
/db\_xref="GABI:939529"  
/db\_xref="taxon:3702"  
/clone="MP1Z20011211Q"  
/issue\_type="root"  
/lab\_host="E. coli TOP10"  
/clone\_id="MP1Z-ADIS-066"  
/note="Vector: PCMVSPORF6; Site 1: SalI; Site 2: NotI;  
cDNA library from Arabidopsis thaliana, accession  
Maslelewiekijs-0; roots from three weeks old plants grown  
on MS-plates at 26M-0C with 16 hours light/day; library  
was made at the Max-Planck-Institute for Plant Breeding  
Research, Cologne, Germany; cloning sites SalI-NotI,  
primer sites and orientation:  
SP6-Sali-CCACGCGCG-5prime-cDNA-polyA-CC-NotI-T7; GATEWAY  
compatible; Note: Sequencing granted in the context of the  
GABI Arabidopsis Verbund I; Genetic Diversity,  
'establishment of high-efficiency SNP-based mapping tools  
and development of methods for genome-wide mutation  
detection'; PI: Bernd Weisshaar Sequence submission managed  
by RZPD/GABI-Primary database: <http://gabi.rzpd.de> This  
clone is available from RZPD; contact RZPD (clone@rzpd.de)  
for further information."

ORIGIN  
Query Match 72.5%; Score 892.2; DB 7; Length 910;  
Best Local Similarity 99.0%; Pred. No. 4,3e-222;  
Matches 897; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 21 AAGTAGAAGAAAATGGCGACTCTTAAGCTTTCTGATTCTGCTCTGCTCTTCTGATGA 80

Db 5 AAGTAGAAGAAAATGGCGACTCTTAAGCTTTCTGATTCTGCTCTGCTCTTCTGATGA 64

Qy 81 TGGTAGCAATTTAGAACCGCTTTTGAAGGATGGGGTACGAAGAGACTTGAATCATATC 140

Db 65 TGGTAGCAATTTAGAACCGCTTTTGAAGGATGGGGTACGAAGAGACTTGAATCATATC 124

Qy 141 AATCTTGGCTCAGAGAGTCTGAACAGAGAAAGTCATGAGCAAGCATACGAAAC 200

Db 125 AATCTTGGCTCAGAGAGTCTGAACAGAGAAAGTCATGAGCAAGCATACGAAAC 184

Qy 201 CTACGCGAAGACCTTCTCAAGACTCTTGACAAAGAGCTCTTAAGATTTGAGAGAGC 260

Db 185 CTACGCGAAGACCTTCTCAAGACTCTTGACAAAGAGCTCTTAAGATTTGAGAGAGC 244

Qy 261 TATCTTGGTGTGACTCTTGAACCGGATGAGGCTTATGAGCTTAATGAGCTAC 320

Db 245 TATCTTGGTGTGACTCTTGAACCGGATGAGGCTTATGAGCTTAATGAGCTAC 304

Qy 321 AAAAAGATGACTTCAAGCAACCAAGTTCTTATGAAGTTGTTGCAAGACATCAAC 380

Db 305 AAAAAGATGACTTCAAGCAACCAAGTTCTTATGAAGTTGTTGCAAGACATCAAC 364

Qy 381 GCACTGCTTCAAGCTTGAAGCACTTCAATGCTGCTTCAAGAGTCTTGAAGAGA 440

Db 365 GCACTGCTTCAAGCTTGAAGCACTTCAATGCTGCTTCAAGAGTCTTGAAGAGA 424

Qy 441 CGTTGCTACCACTACCGGTGACTTCAAGAAAGCTTTGTTGTTCTTGTATACCTAT 500

Db 425 CGTTGCTACCACTACCGGTGACTTCAAGAAAGCTTTGTTGTTCTTGTATACCTAT 484

Qy 501 CAGGTACGAAGAGATGAAGTGAATGATGATGCTTGAAGCAAGAGCTTACGTGCA 560

Db 485 CAGGTACGAAGAGATGAAGTGAATGATGATGCTTGAAGCAAGAGCTTACGTGCA 544

Qy 561 TGAAGAAATCAAGAGCAAGCACTCAATGATGAGATGTTATTAATCTTGTCCACAG 620

Db 545 TGAAGAAATCAAGAGCAAGCACTCAATGATGAGATGTTATTAATCTTGTCCACAG 604

Qy 621 AACCAAGCTCAGATGAATGCTACTTTAAAGCTTCAAGAGATGATGAGAGAAAT 680

Db 605 AACCAAGCTCAGATGAATGCTACTTTAAAGCTTCAAGAGATGATGAGAGAAAT 664

Qy 681 TCTCAAGAGCTTGAAGAGAGATGATGATGATGATGATGATGATGATGATGATGAT 740

Db 665 TCTCAAGAGCTTGAAGAGAGATGATGATGATGATGATGATGATGATGATGATGAT 724

Qy 741 CATTCAGTCTTGAAG 800

Db 725 CATTCAGTCTTGAAG 784

Qy 801 CAAATCTGAATCAAG 860

Db 785 CAAATCTGAATCAAG 844

Qy 861 CTTGAAGCTTGAAG 920

Db 845 CTTGAAGCTTGAAG 904

Qy 921 TACCAA 926

Db 905 TACCAA 910

RESULT 8  
LOCUS CENSOC11  
DEFINITION CENSOC11 865 bp mRNA linear HTC 06-FEB-2004  
Arabidopsis thaliana Full-length cDNA Complete sequence from clone  
GSLRFB692P06 of Flowers and buds of strain col-0 of Arabidopsis  
thaliana (thale cress).  
ACCESSION EX814350

```

VERSION      BX814350.1  GI:42472041
KEYWORDS     HTC; GSUT cDNA.
SOURCE       Arabidopsis thaliana (thale cress)
ORGANISM     Arabidopsis thaliana
              Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
              Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
              rosids; eustosids II; Brassicales; Brassicaceae; Arabidopsids.
REFERENCE    1 (bases 1 to 865)
AUTHORS      Castelii,V., Aury,J.M., Jallion,O., Wincker,P., Clepet,C.,
              Menard,M., Cruaud,C., Quetier,F., Scarpelli,C., Schacher,V.,
              Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.
TITLE        Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences:
              A Combined Approach to Evaluate and Improve Arabidopsis Genome
              Annotation
JOURNAL      Unpublished
REFERENCE    2 (bases 1 to 865)
AUTHORS      Genoscope.
TITLE        Direct Submission
JOURNAL      Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
              BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
              - Web : www.genoscope.cns.fr)
COMMENT      The sequences are based on single pass reads.
              Life Technologies (a division of Invitrogen) members carried out
              full-length libraries construction : Temple G.
              Genoscope members carried out sequencing and annotation : Castelii
              V., Aury J.M., Jallion O., Wincker P., Menard M., Cruaud C.,
              Schacher V., Weissenbach J., Salanoubat M.
              URGV INRA : Clepet C., Caboche M.
              Annotation is based on the June 2003 version of the Arabidopsis
              genome released by MIPS (Munich Information center for Protein
              Sequences) . 5 prime and 3 prime are assembled with Phrap.
              http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full
              length
              http://www.genoscope.cns.fr/cgi-bin/ggb?source=Arabidopsis.
FEATURES
  source
    location/Qualifiers
      1..865
      /organism="Arabidopsis thaliana"
      /mol_type="mRNA"
      /strain="Col-0"
      /db_xref="taxon:3702"
      /clone="GSUTRB692F06"
      /tissue_type="Flowers and buds"
      /plasmid="PCMVSPORT_6"
      1..865
      /gene="At1g35720"
ORIGIN
Query Match      68.0%; Score 836.2; DB 3; Length 865;
Best Local Similarity 97.9%; Pred. No. 1.9e-207;
Matches 847; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 349 CTTATGGAAGTTCCTGCACAAGACATCAACGACGTGTTACCGCTAGCAAGCTTAC 408
DB 1 CTTATGGAAGTTCCTGCACAAGACATCAACGACGTGTTACCGCTAGCAAGCTTAC 60
QY 409 CATGCTCGCTACAAGAAGTCTCTTGAAGAAGAGTGTGCTACCACTACCGGTGACTTC 468
DB 61 CATGCTCGCTACAAGAAGTCTCTTGAAGAAGAGTGTGCTACCACTACCGGTGACTTC 120
QY 469 AGAAAGCTTTGGTTCTCTTGTACCTCATACAGGTACGAGAGATGAATGAACATG 528
DB 121 AGAAAGCTTTGGTTCTCTTGTACCTCATACAGGTACGAGAGATGAATGAACATG 180
QY 529 ACATTGGCTTACGAAGAAGTACGTGTCATGAGAAAATCAAGACAAAGCACTACAT 588
DB 181 ACATTGGCTTACGAAGAAGTACGTGTCATGAGAAAATCAAGACAAAGCACTACAT 240
QY 589 GATGAGAGATGTTATTAATCTGTCCACAAGAAGCAAGCTCAGATCATGCTACTTTT 648
DB 241 GATGAGAGATGTTATTAATCTGTCCACAAGAAGCAAGCTCAGATCATGCTACTTTT 300
QY 649 AACCGTTTACCAAGATGATCATGCGCAGGAATTTCTCAAGAGTCTTGAAGAGAGATGAT 708

```

```

DB 301 AACCGTTTACCAAGATGATCATGCGCAGGAATTTCTCAAGAGTCTTGAAGAGAGATGAT 360
QY 709 GATGACAAAGTTCCTTGCACTTTTGAAGGTCAACCATTCAGTGTGTCACAAGACAGCTT 768
DB 361 GATGACAAAGTTCCTTGCACTTTTGAAGGTCAACCATTCAGTGTGTCACAAGACAGCTT 420
QY 769 TACTTGTGATGTTCTTGTGTTTGTGTCACATCAACAAATGTAAGTATGAAGACACTC 828
DB 421 TACTTGTGATGTTCTTGTGTTTGTGTCACATCAACAAATGTAAGTATGAAGACACTC 480
QY 829 ACTAGAAATTTGATCCACAAGAGCTGAGATTGCTTGAAGGTATTTGAGAGAGATTCAG 888
DB 481 ACTAGAAATTTGATCCACAAGAGCTGAGATTGCTTGAAGGTATTTGAGAGAGATTCAG 540
QY 889 CGCAGGAACAGACTTCCTTGTGAGAAAGCTATTACCAAGACACTCGTGGATTAACGAG 948
DB 541 TGCAGGAACAGACTTCCTTGTGAGAAAGCTATTACCAAGACACTCGTGGATTAACGAG 600
QY 949 AAGATGCTGTGCGCACTTCTCGGTGAAAGATGCTTAATCAATCAATCTCCACAGAGA 1008
DB 601 AAGATGCTGTGCGCACTTCTCGGTGAAAGATGCTTAATCAATCAATCTCCACAGAGA 660
QY 1009 AACATTAAGTGTCTTACAGACTTCTGTATCTTTATCTTCCCTCTCTCTTTGATGAG 1068
DB 661 AACATTAAGTGTCTTACAGACTTCTGTATCTTTATCTTCCCTCTCTCTTTGATGAG 720
QY 1069 TTTCAAAATGTTGATTTGTTGTTTCTACAAAACCTGTTGTTGTTGTTGTTGTTGA 1128
DB 721 TTTCAAAATGTTGATTTGTTGTTTCTACAAAACCTGTTGTTGTTGTTGTTGTTGA 780
QY 1129 GTTCTTAATTAATGCAAAAAGAGAGACAGAGAACAGTGTGCTCTTAAGTTAAT 1188
DB 781 GTTCTTAATTAATGCAAAAAGAGAGAACAGTGTGCTCTTAAGTTAAT 840
QY 1189 ATATATGAAGACATTTGGCTTAAA 1213
DB 841 ATATATGAAGACATTTGGCTTAAA 865

RESULT 9
LOCUS       CF652048                      843 bp      mRNA      linear      EST 06-NOV-2003
DEFINITION  30-1020580-066-004-L08-SP6P MP1Z-ADIS-066 Arabidopsis thaliana cDNA
            Clone MP1Zp2001L084Q 5-PRIME, mRNA sequence.
ACCESSION   CF652048
VERSION     CF652048.1  GI:37428171
KEYWORDS    EST
SOURCE      Arabidopsis thaliana (thale cress)
ORGANISM    Arabidopsis thaliana
              Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
              Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
              rosids; eustosids II; Brassicales; Brassicaceae; Arabidopsids.
REFERENCE    1 (bases 1 to 843)
AUTHORS      Schmidt,K.U., Soerensen,T.R., Stracke,R., Torjek,O., Altmann,T.,
              Mitchell-Olds,T. and Weisshaar,B.
TITLE        Large-scale identification and analysis of genome-wide
              single-nucleotide polymorphisms for mapping in Arabidopsis thaliana
JOURNAL      Genome Res. 13 (6), 1250-1257 (2003)
MEDLINE     22683290
PUBMED     12799357
COMMENT      Contact: Weisshaar B
              ADIS DNA core facility at MP1Z
              Max-Planck-Institute for Plant Breeding Research
              Carl-von-Linne Weg 10, 50829 Koeln, Germany
              Fax: 00492215062851
              Email: weisshaar@mp1z-koeln.mpg.de
              Insert Length: 843 Std Error: 0.00
              Plate: 4 row: L column: 08
              Seq primer: SP6P.
FEATURES
  source
    location/Qualifiers
      1..843
      /organism="Arabidopsis thaliana"
      /mol_type="mRNA"

```

```

/ecotype="Ms-0"
/db_xref="GABI:938637"
/db_xref="taxon:3703"
/clone="MP12p2001L084Q"
/tissue_type="root"
/lab_host="E. coli TOP10"
/clone_1fb="MB12-ADIS-066"
/notes="vector: pGMWArabid6; Site 1: SalI; site 2: NotI; cDNA library from Arabidopsis thaliana, accession Maslowski.kja-0; roots from three weeks old plants grown on MS-plates at 26m-0°C with 16 hours light/day; library was made at the Max-Planck-Institute for Plant Breeding Research, Cologne, Germany; cloning sites SalI-NotI, primer sites and orientation: SP6-SalI-CCGCGCCGCG-5prime-CDNA-polyA-CC-NotI-T7; GATEWAY compatible; Note: Sequencing granted in the context of the GABI Arabidopsis Verbund I: Genetic diversity, 'Establishment of high-efficiency SNP-based mapping and development of methods for genome-wide mutation detection', PI: Bernd Weisshaar Sequence submission managed by RZPD/GABI-Primary database: http://gabi.rzpd.de This clone is available from RZPD; contact RZPD (clone@rzpd.de) for further information."
```

**ORIGIN**

Query Match	67.6%	Score 831.8	DB 7	Length 843
Best Local Similarity	99.2%	Pred. No. 2.7e-206		
Matches 836	Conservative 0	Mismatches 7	Indels 0	Gaps 0
QY	21	AAAGTAAAGAAAAATGGCGACTCTTAAGGTTCTGATTCGTGCTCGCTCTCTGATGA	80	
Db	1	AAAGTAAAGAAAAATGGCGACTCTTAAGGTTCTGATTCGTGCTCGCTCTCTGATGA	60	
QY	81	TGCTGAGCAATTGGAACCCGCTTTTGAAGATGGGGTTCGAACGAGAGACTTGATCATTC	140	
Db	61	TGCTGAGCAATTGGAACCCGCTTTTGAAGATGGGGTTCGAACGAGAGACTTGATCATATC	120	
QY	141	AATCTTGGCTCACAAGAAAGTGTGCAACAGAGAAAGTCATCAGGCAAGCATACAGAAAC	200	
Db	121	AATCTTGGCTCACAAGAAAGTGTGCAACAGAGAAAGTCATCAGGCAAGCATACATGAAC	180	
QY	201	CTACGGGAGACCTTCTCAAGACTCTTGAACAAGAGCTCTTACGATTTGAGAGAC	260	
Db	181	CTACGGGAGAACCTTCTCAAGACTCTTGAACAAGAGCTCTTACGACTTCGAGAGAC	240	
QY	261	TATCTTGTGTGAGCTCTTGAAACCCGGAGACGTGATGCTTATTTGGTATGAAAGTAC	320	
Db	241	TATCTTGTGTGAGCTCTTGAAACCCGGAGACGTGATGCTTATTTGGTATGAAAGTAC	300	
QY	321	AAAAAGATGACCTTCAAGCAACCAAGTCTTATGGAAGTTGCTTGACAAGAGCATCAAC	380	
Db	301	AAAAAGATGAGCTTCAAGCAACCAAGTCTTATGGAAGTTGCTTGACAAGAGCATCAAC	360	
QY	381	GCAGCTGCTTCAGGCTTGGCGAAGTTCACAGTCGCTACAGAAAGTCTTGAAGAGGA	440	
Db	361	GCAGCTGCTTCAGGCTTGGCGAAGTTCACAGTCGCTACAGAAAGTCTTGAAGAGGA	420	
QY	441	CGTTGCTCACCACTACCCGGTGACTTCAGAAAGCTTTGGTCTCTGTGTACCTATA	500	
Db	421	CGTTGCTCACCACTACCCGGTGACTTCAGAAAGCTTTGGTCTCTGTGTACCTATA	480	
QY	501	CAGGTACGAAGGAGATGAATGAACTATGACATTTGGCTTAAGCAAGAGCTAAGCTGTCCA	560	
Db	481	CAGGTACGAAGGAGATGAATGAACTATGACATTTGGCTTAAGCAAGAGCTAAGCTGTCCA	540	
QY	561	TGAGAAATATCAAGAGACACTACAAATGATGAGAGATGTTATTTGAATCTTGTCACAAAG	620	
Db	541	TGAGAAATATCAAGAGACACTACAAATGATGAGAGATGTTATTTGAATCTTGTCACAAAG	600	
QY	621	AAGCAAAAGCTCAGATCAATGCTACTTTTAAACCGTTACCAAGATGATCATGGCGAGAAAT	680	
Db	601	AAGCAAAAGCTCAGATCAATGCTACTTTTAAACCGTTACCAAGATGATCATGGCGAGAAAT	660	

Oy		681	CTCAGAGGCTTTAGAGAAGAGATGTATGATGAACAATTCTTGCACCTTTTAGGCTCAAC	740
Dd		661	TCTCAAGAGCTTTAGAGAAAGAGATGATGATGACAAAGTTCCTTGCTTTGGAGGCTAAC	720
Oy		741	CATTAGAGCTTTGACAAGACAGAGCTTTACTTTGTGATGTGTTCTTGTCAGCATCA	800
Dd		721	CATTAGAGCTTTGACAAGACAGAGCTTTACTTTGTGATGTGTTCTTGTCAGCATCA	780
Oy		801	CAAACTTGGAAGCTGATGAGAGGACCTCATTAAGATTGTGACCACAAGAGCTGAGATTGA	860
Dd		781	CAAACTTGGAAGCTGATGAGAGGACCTCATTAAGATTGTGACCACAAGAGCTGAGATTGA	840
Oy		861	CTT 863	
Dd		841	CTT 843	
	RESULT 10			
	LOCUS	CA782049	814 bp	mRNA linear EST 03-DEC-2002
	DEFINITION	01E11AIF Infected Arabidopsis Leaf Arabidopsis thaliana cDNA, mRNA sequence.		
	ACCESSION	CA782049		
	VERSION	CA782049.1	GI:26020095	
	KEYWORDS	EST.		
	SOURCE	Arabidopsis thaliana (thale cress)		
	ORGANISM	Arabidopsis thaliana		
	REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.		
	AUTHORS	I (bases 1 to 814) Lundsgaard,M., Emmersen,J., Nielsen,K.L., Wilson,I., Somerville,S. and Weinder,K.G.		
	TITLE	EST sequencing of Erysiphe cichoracearum infected Arabidopsis plants		
	JOURNAL	Unpublished (2002)		
	COMMENT	Contact: Karen G. Welinder Institut for bioteknologi Aalborg Universitet Sohnegaardsholmsvej 49, 9000 Aalborg, Denmark Tel.: +45 96358467 Fax: +45 98141808 Email: kgw@bio.auc.dk.		
	FEATURES	location/Qualifiers		
	source	1..814		
		/organism="Arabidopsis thaliana"		
		/mol_type="mRNA"		
		/ecotype="Columbia"		
		/db_xref="taxon:3702"		
		/dev_stage="Plant 3 weeks old, three days post infection"		
		/clone_lib="Infected Arabidopsis leaf"		
		/note="Organ: leaf; Vector: pBluescript; Mixed cDNA library of Arabidopsis and E. cichoracearum infected leaf from three weeks old Arabidopsis plants. Plants were harvested 3 days after infection and mRNA oligo dt selected."		
	ORIGIN			
	Query Match	63.9%;	Score 786.2;	DB 6; Length 814;
	Best Local Similarity	99.4%;	Pred. No. 2.3e-194;	
	Matches 810;	Conservative 0;	Mismatches 3;	Indels 2; Gaps 2;
Oy		28	AGAAAAATGGCACTCTTAAGGTTTTGATTGATTCGTGCTCTCTTGTGATGATGGCGAG	87
Dd		1	AGAAAATGGCACTCTT -AGTTTTCTGATTCGTGCTCTCTTGTGATGATGGCTGAG	59
Oy		88	CAATTGGAACCGCTTTGAAGATGGGATACGAACGAGACTTGATCATATCAATCTTG	147
Dd		60	CAATTGGAACCGCTTTGAAGATGGGATACGAACGAGACTTGATCATATCAATCTTG	119
Oy		148	GCTTACGAAAGTGTGTAACGAGAAAGTATATAGGACAGATACCAAGAACTTACGGC	207
Dd		120	GCTTACGAAAGTGTGTAACGAGAAAGTATATAGGACAGATACCAAGAACTTACGGC	179



LOCUS BU636350 775 bp mRNA linear EST 23-SEP-2002  
 DEFINITION 049H05 Infected Arabidopsis leaf Arabidopsis thaliana cDNA, mRNA  
 ACCESSION BU636350  
 VERSION BU636350.1 GI:23303605  
 KEYWORDS EST.  
 SOURCE Arabidopsis thaliana (thale cress)  
 ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 1 (bases 1 to 775)  
 Lundgaard, M., Emmersen, J., Nielsen, K.L., Wilson, I., Somerville, S. and Wellinger, K. G.  
 EST sequencing of Erysiphe cichoracearum infected Arabidopsis plants  
 JOURNAL Unpublished (2002)  
 COMMENT Contact: Karen G. Wellinger  
 Institut for bioteknologi  
 Aalborg Universitet  
 Sohngaardsholmsvej 49, 9000 Aalborg, Denmark  
 Tel: +45 96358467  
 Fax: +45 98141808  
 Email: kgw@bio.auc.dk.  
 FEATURES  
 source location/Qualifiers  
 1..775  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /ecotype="Columbia"  
 /db\_xref="taxon:3702"  
 /dev\_stage="plant 3 weeks old, three days post infection"  
 /clone\_lib="infected Arabidopsis leaf"  
 /note="Organ: Leaf; Vector: pluscript; Mixed cDNA library of Arabidopsis and E. cichoracearum infected leaf from three weeks old Arabidopsis plants. Plants were harvested 3 days after infection and mRNA oligo dt selected."

ORIGIN  
 Query Match 59.7%; Score 734.4; DB 5; Length 775;  
 Best Local Similarity 99.5%; Pred. No. 8e-181;  
 Matches 768; Conservative 0; Mismatches 1; Indels 3; Gaps 3;  
 21 AAGTAGAAGAAAATGCGACTCTTAAGGTTTGGATTCTGCTCTCTTGTATGA 80  
 6 AGGTAGAGAAAATGCGACTCTTAAGG-TTCTGATTCTG-TCTCTCTCTCTGTATGA 63  
 81 TGTGAGCAATTGAGAACCGCTTTTGAAGATGGGGTACGACGAGGACTTGATCATATC 140  
 64 TGTGTAGCAATTGAGAACCGCTTTTGAAGATGGGGTACGACGAGGACTTGATCATATC 123  
 141 AATCTTGGCTACAGAAAGTCTGAACAGAGAAAAGTCATGAGGCAAGCATACGAGAAC 200  
 124 AATCTTGGCTACAGAAAGTCTGAACAGAGAAAAGTCATGAGGCAAGCATACGAGAAC 183  
 201 CTACGGGGAAGACTTTTCAAGACTCTTGACAGAGAGCTCTTACGATTTGAGAGAGC 260  
 184 CTACGGGGAAGACTTTTCAAGACTCTTGACAGAGAGCTCTTACGATTTGAGAGAGC 243  
 261 TATCTTGTGTGAGACTCTTAACCCGGTGAAGCGGTATTTATTTGGCTAATGAAGTAC 320  
 244 TATCTTGTGTGAGACTCTTAACCCGGTGAAGCGGTATTTATTTGGCTAATGAAGTAC 303  
 321 AAAAAGATGAGACTTCAAGCAACCAAGTTCTTATGGAAGTGTCTTGACAAAGACATCAAC 380  
 304 AAAAAGATGAGACTTCAAGCAACCAAGTTCTTATGGAAGTGTCTTGACAAAGACATCAAC 363  
 381 GCGAGTCTTCAAGCTTACGCAAGGTTACATGCTCGTTACAGAAAGTCTTTGAAGGGA 440  
 364 GCGAGTCTTCAAGCTTACGCAAGGTTACATGCTCGTTACAGAAAGTCTTTGAAGGGA 423  
 441 CGTTGCTCACCACTACGCGGTGACTTCAAGAAAGCTTTTGTTTCTTGTACCTCATAT 500  
 |||||

Db 424 CGTTGCTCACCACTACCGGTGACTTCAAGAAAGCTTTTGTTCTTGTACCTCATAT 483  
 Qy 501 CAGGTAGAAAGAGATGAATGAACATGACATTTGGCTTAAGCAAGAAAGTGTGTCA 560  
 Db 484 CAGGTAGAAAGAGATGAATGAACATGACATTTGGCTTAAGCAAGAAAGTGTGTCA 543  
 Qy 561 TGAGAAAATCAAGACAAAGACATACATGATGATGAGATGATTAATGAATCTTGTCCCAAG 620  
 Db 544 TGAGAAAATCAAGACAAAGACATACATGATGATGAGATGATTAATGAATCTTGTCCCAAG 603  
 Qy 621 AAGCAAGCTCAGATCAATCTCTTTTAACCGTTACCAAGATGATGATGAGGAGAAAT 680  
 Db 604 AAGCAAGCTCAGATCAATCTCTTTTAACCGTTACCAAGATGATGATGAGGAGAAAT 663  
 Qy 681 TCTCAAGAGTCTTGAAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 719  
 Db 664 TCTCAAGAGTCTTGAAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 723  
 Qy 740 CCAATCAAGCTTGAAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 791  
 Db 724 CCAATCAAGCTTGAAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 775  
 RESULT 13  
 AV784024/c 703 bp mRNA linear EST 28-MAR-2002  
 LOCUS AV784024 RAFL5 Arabidopsis thaliana cDNA clone RAF105-15-G22 3',  
 DEFINITION mRNA sequence.  
 ACCESSION AV784024  
 VERSION AV784024.1 GI:19802814  
 KEYWORDS EST.  
 SOURCE Arabidopsis thaliana (thale cress)  
 ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 1 (bases 1 to 703)  
 Seki, M., Narusaka, M., Ishida, J., Kamiya, A., Satou, M., Nakajima, M., Oono, Y., Sakurai, T., Carninci, P., Kawai, J., Itoh, M., Ishii, Y., Arakawa, T., Shibata, K., Shinagawa, A., Muramatsu, M., Hayashizaki, Y. and Shinozaki, K.  
 Large scale analysis of Arabidopsis full-length cDNA (2002b)  
 JOURNAL Unpublished (2002)  
 COMMENT Contact: Motoaki Seki  
 Plant Functional Genomics Research Group  
 RIKEN Genomic Sciences Center  
 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan  
 Tel: 81-298-36-4359  
 Fax: 81-298-36-9060  
 Email: mseki@rvc.riken.go.jp  
 An Arabidopsis full-length cDNA library was constructed essentially as reported previously (Seki et al., 1998). This clone is in a modified pluscript vector as a Set1/XhoI insert. Please visit our web site ([http://www.gsc.riken.go.jp/e/plant/index\\_e.html](http://www.gsc.riken.go.jp/e/plant/index_e.html)) for further details.  
 FEATURES  
 source location/Qualifiers  
 1..703  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3702"  
 /clone="RAFL05-15-G22"  
 /dev\_stage="rossette plants"  
 /lab\_host="SOLR"  
 /clone\_lib="RAFL5"  
 /note="Site 1: Set1; Site 2: XhoI; subjected to dehydration-treated(1,2,5,10,24 hr)"

ORIGIN  
 Query Match 57.2%; Score 703; DB 1; Length 703;  
 Best Local Similarity 100.0%; Pred. No. 1.3e-172;  
 Matches 703; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 507 CGAAGAGATGAAGTGAACATGATGCTTAAGCAAGAAAGTGAAGTGTGATGAGAA 566  
 |||||







**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model1

Run on: August 21, 2005, 22:23:23 ; Search time 252 Seconds

(without alignments)  
7986.589 Million cell updates/sec

Title: US-10-690-564-1

Perfect score: 1230  
Sequence: 1 cccgcgcctccgaacactaa.....aaaaaaaaaaaaaaaaaa 1230

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Issued Patents NA:  
1: /cgn2\_6/prodata/1/ina/5A.COMB.seq:\*  
2: /cgn2\_6/prodata/1/ina/5B.COMB.seq:\*  
3: /cgn2\_6/prodata/1/ina/6A.COMB.seq:\*  
4: /cgn2\_6/prodata/1/ina/6B.COMB.seq:\*  
5: /cgn2\_6/prodata/1/ina/PCITUS.COMB.seq:\*  
6: /cgn2\_6/prodata/1/ina/backfile1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	468.6	38.1	1293	US-09-325-932A-17	Sequence 17, App1
2	364.2	29.6	789	US-09-325-932A-20	Sequence 20, App1
3	208.4	16.9	704	US-09-325-932A-21	Sequence 21, App1
4	186.2	15.1	484	US-09-325-932A-18	Sequence 18, App1
5	149.2	12.1	1332	US-09-949-016-2081	Sequence 2081, App
6	149.2	12.1	1339	US-09-023-655-1149	Sequence 1149, App
7	149.2	12.1	1339	US-09-949-016-722	Sequence 722, App
8	138.4	11.3	221	US-09-325-932A-19	Sequence 19, App1
9	103.8	8.4	1362	US-09-023-655-1152	Sequence 1152, App
10	103.8	8.4	1362	US-09-814-915A-101	Sequence 101, App
11	103.8	8.4	1614	US-09-949-016-2431	Sequence 2431, App
12	102.8	8.4	1660	US-09-919-039-393	Sequence 393, App
13	101.6	8.3	981	US-09-324-096A-1	Sequence 1, App1
14	101.6	8.3	981	US-09-324-096A-3	Sequence 3, App1
15	101.6	8.3	981	US-09-324-096A-5	Sequence 5, App1
16	101.6	8.3	981	US-09-970-969-1	Sequence 1, App1
17	101.6	8.3	981	US-09-970-969-3	Sequence 3, App1
18	101.6	8.3	981	US-09-970-969-5	Sequence 5, App1
19	100	8.1	957	US-08-125-746-4	Sequence 4, App1
20	100	8.1	1460	5225537-3	Patent No. 5225537
21	100	8.1	1460	5225537-3	Patent No. 5225537
22	100	8.1	1567	US-08-125-746-2	Sequence 2, App1
23	100	8.1	1632	US-09-949-016-4306	Sequence 4306, App
24	98.4	8.0	1399	US-09-023-655-1342	Sequence 1342, App
25	98	8.0	2305	US-08-526-136-1	Sequence 1, App1
26	98	8.0	2311	US-08-526-136-3	Sequence 3, App1
27	96.8	7.9	1399	US-09-949-016-5663	Sequence 5663, App

28	95.2	7.7	1301	4	US-09-023-655-1190	Sequence 1190, App
29	95.2	7.7	1516	4	US-09-919-172-38	Sequence 38, App1
30	93.6	7.6	1957	4	US-09-949-016-4660	Sequence 4660, App
31	92.2	7.5	1932	4	US-09-949-016-1750	Sequence 1750, App
32	85.8	7.0	1300	4	US-09-949-016-2753	Sequence 2753, App
33	85.8	7.0	5300	4	US-09-949-016-14495	Sequence 14495, App
34	81.4	6.6	2054	4	US-09-949-016-1199	Sequence 1199, App
35	81.4	6.6	2054	4	US-09-949-016-1200	Sequence 1200, App
36	81.4	6.6	2120	4	US-09-949-016-1197	Sequence 1197, App
37	81.4	6.6	2120	4	US-09-949-016-1198	Sequence 1198, App
38	81.4	6.6	2137	4	US-09-976-594-348	Sequence 348, App
39	80.8	6.6	2386	4	US-09-949-016-4604	Sequence 4604, App
40	80.8	6.6	2386	4	US-09-949-016-4605	Sequence 4605, App
41	80	6.5	1356	4	US-09-010-147B-15	Sequence 15, App1
42	64.8	5.3	2105	4	US-09-799-451-493	Sequence 493, App
43	63	5.1	1432	4	US-09-799-451-269	Sequence 269, App
44	59.4	4.8	917	4	US-09-799-451-268	Sequence 268, App
45	59.4	4.8	1211	4	US-09-799-451-270	Sequence 270, App

## ALIGNMENTS

```
RESULT 1
US-09-325-932A-17
; Sequence 17, Application US/09325932A
; Patent No. 6451604
; GENERAL INFORMATION:
; APPLICANT: Flinn, Barry
; APPLICANT: Lasham, Annette
; TITLE OF INVENTION: Compositions affecting programmed cell
; FILE REFERENCE: 1022
; CURRENT APPLICATION NUMBER: US/09/325,932A
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 206
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 17
; LENGTH: 1293
; TYPE: DNA
; ORGANISM: Pinus radiata
US-09-325-932A-17

Query Match      38.1%; Score 468.6; DB 3; Length 1293;
Best Local Similarity 68.4%; Pred. No. 1.3e-127;
Matches 664; Conservative 0; Mismatches 304; Indels 3; Gaps 1;
```

QY	30	AAAAATGGGACCTTAAAGGTTCTGATTCCTGCTCCTTGATGATGCTGAGCA	89
DB	91	AACAAATGTGACTCTACCGTCCCGACGACCTGCCCCCTGATGCGATGCTGCGAGCA	150
QY	90	ATTGAGAACCGCTTTTGAAGGATGAGGTACGAACTTGATCATATCAATCTTGGC	149
DB	151	GCTCCGAGACGCTTGCGAGGATGAGGAACTAATGAAAGCTGATCATATTTGGG	210
QY	150	TCCAGAAAGTCTGAAACAGAGAAATCATACGACGATACCAAGAACTTACGCGCA	209
DB	211	TCTATGAAATGCGCGCCAGAGGAACTGATTCGGCAAACTTATCGGACCTTACGCGCA	270
QY	210	AGACCTTTCAGACCTTGAAGAGGCTCTTAAGATTTTGAGAGAGCTATCTTGGT	269
DB	271	GGACTCTCTCAAGCAATTTGACAGAACTTACCAATGATTTGAGAGGCTGTGTCTCT	330
QY	270	GTTGACTTTGAACCCGATGAGGCTGATCTTATTTGCTTAATGAAGTACAAAGATG	329
DB	331	TTTGTCATCTGATTCGGCTGGAAGTATGCTGATCTTGGGATGAGGACGAAAGATG	390
QY	330	GACTTCAAGCAACAGTTCTTATGGAAGTCTTGCACAGAGCATCAACGAGCTGCT	389
DB	391	GACTTCAACCAACAGGTTCTCATGGAATTAATCTGACAGAGGTCCTCCACAGTTGCT	450
QY	390	TACGCTAGGCAAGCTTACATGCTGCTCAAGAAAGTCTTGAAGAGAGAGCTGCTCA	449

Db 451 TATGCAAGACAAGCATATCATGCCGATACAAAGATCAATGAAAGAGGACCTCGCTCA 510  
QY 450 CCACACTACCGGTGACTTTCAGAAAGCTTTGGTTTCTCTGTTAACTTCATACAGATACGA 509  
Db 511 CCACACAACCTGGAGATTTCGTAAAGTTCGTGACCTCTTGGAGGCTCTCAACGTATGA 570  
QY 510 AGGAGATGAAGTGAACATGACATTTGGCTAAGCAAGAGCTAACGTGGTCCATGAGAAAT 569  
Db 571 TGGAGATGAGTGAATATGACTTTGGCAAAAGCAGAGGCTAAGATACTCCACAGAAAGAT 630  
QY 570 CAAGGACAGCACTACATATGATGAGATGTTATTAATCTTGTCCAAGAAAGCAAGC 629  
Db 631 CTGAGAGAAAGCTTATGGCCATGAGATCTCATAGGATTTGGCTACTAGAGCAAGC 690  
QY 630 TCAGATCAATGCTACTTTTAACGCTTACCAAGATGATCAATGAGGAGAAATTCGAAG 689  
Db 691 ACGAGTCAATGCTACGCTGATCACTACAAAGATGATTGAAATGATATCAAGAGA 750  
QY 690 TCTTGAAGAAAGAGATGATGATGACAAATCTTGGCACTTTGAGGTCAACCATGAGT 749  
Db 751 TTT---GAAAGCTGATCCAAAGACGCGTTCCTTACTATCTGAGAGCTACAGTAAAGT 807  
QY 750 CTTGACAAGCCAGAGCTTTACTTTGCTGATGTTCTTCTGTTCAAGCAATCAAGAACTG 809  
Db 808 CCTGACTCGCCCTGAGAGATATTTGAAAGGTTCTTCTGCTAGCCATCAATAGGAGG 867  
QY 810 AACTGATGAAGAGCACTCACTGAATTTGACCAAGAGCTGAGATTGACTTGAAGT 869  
Db 868 AACGATGAAGGAGGCTTGACCAAGATGTTGCTTCCAGGAGCCGAGGTTGATGAAGTT 927  
QY 870 CATTGAGAGAGTACCAAGGACAGAACAGCATTTCTTTGAGAAAGCTATTAACAAGA 929  
Db 928 TATTAAGTGAAGAGTACCAAGAGGAAATAGCATCCCTCGATCGATGCTATGTCAGAGA 987  
QY 930 CATCTGCGAGATTACAGAAAGATCTCTGTCGCACTTCTGCGTGAAGATGATCTTAATC 989  
Db 988 CACTTACTGAGACTATGAAAAATGCTTCTGCGATGATGAGCCACGTCAGAGCTTGATT 1047  
QY 990 AATCAATCTCTC 1000  
Db 1048 TACAGTACTC 1058

RESULT 2  
US-09-325-932A-20  
; Sequence 20, Application US/09325932A  
; Patent No. 6451604  
; GENERAL INFORMATION:  
; APPLICANT: Flinn, Barry  
; APPLICANT: Lasham, Annette  
; TITLE OF INVENTION: Compositions affecting programmed cell  
; TITLE OF INVENTION: death and their use in the modification of forestry plant develo  
; FILE REFERENCE: 1022  
; CURRENT APPLICATION NUMBER: US/09/325,932A  
; NUMBER OF SEQ ID NOS: 206  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 20  
; LENGTH: 789  
; TYPE: DNA  
; ORGANISM: Pinus radiata  
US-09-325-932A-20

Query Match 29.6%; Score 364.2; DB 3; Length 789;  
Best Local Similarity 69.7%; Pred. No. 6,1e-97;  
Matches 508; Conservative 0; Mismatches 218; Indels 3; Gaps 1;

QY 25 AGAAGAAAATGCGACTTAAAGTTTCTGATTCCTGCTCTCTCTGATGATGCT 84  
Db 64 AGAGAGAGATGCGACTATCGCGTCCACCGCTGCTCTCTCGGCTGAGATGCC 123  
QY 85 GAGCAATTGAAACCGCTTTTGAAGATGGGTTACGAACAGAGACTTGATCATATCATC 144

Db 124 GAGCAGCTCCAAAAAGCTTTCCAGAGATGGGGAGCAATGAAGATCTGATCATATCCATA 183  
QY 145 TTGGCTCAGAGAAGCTTGAACAAGAGAAATGATCAAGGCAAGCATACAGAAACCTTC 204  
Db 184 CTGCTTCAGAGAAACGACAGCGAGAAAGTAAATCCGAAACATATGCCCAGACATAT 243  
QY 205 GCGAAGACCTTCTCAAGACTTTTGAACAAGAGCTCTCAAGATTTCCAGAGAGCTATC 264  
Db 244 GGGGAAGATCTTCTCAAGCGCTTGAACAAGAACTCTCTAAGTACTTTGAGAGATCTG 303  
QY 265 TTGTTGAGACTCTTGAACCCGATGAGCTGATGCTTTATGAGCTTAATGAGCTACAAA 324  
Db 304 CTTCGTGAGACCTGATCTCGAGAGCTGATGATCTTGTCCATGAACCTACCAAG 363  
QY 325 AGATGAGCTTCAAGCAACCAATTTCTTAAGAGTTGCTTGAACAAGCATCAACGAG 384  
Db 364 AGATGAGCTTCAAGCAACCTGAGTTCTCAAGAAATTTCTTGAACAAGCTTCAATGAG 423  
QY 385 CTGCTTCAAGCTAAGGCACTTACATGCTCGCTCAAGAAAGCTCTTGAAGAGAGCTT 444  
Db 424 TATTCATGATGAGGAGGCTATCATGCTCTGTTAAGAAATCTTGAAGAGAGATC 483  
QY 445 GCTCACACACTACCGGTGACTTCAAGAAAGCTTTGTTTCTTCTGTTAATCTCATACAG 504  
Db 484 GATATCACACTACGAGGATTTCCGCAAGCTGCTTGTCTCTGCAAGTACCTTTCGG 543  
QY 505 TCGAAGAGATGAAGTGAACATGACATTTGGCTTAAGCAAGAGCTAAGCTGATGAG 564  
Db 544 TATGAGGAGCTGAGTGAACATGATTTGGAGATCAGAGGCTTAAGATCTTCAATGAG 603  
QY 565 AAAATCAAGACAGCACTACATGATGAGATGTTATTAAGATCTTGTCCCAAGAAAGC 624  
Db 604 AGATTCAGAGAGAGCTTACATCATATGATGAGCTCATGAAATGTTACTCAAGAAAGT 663  
QY 625 AAAAGCTCAGATCAATGCTTCTTAAACCTTACCAAGATGATCAAGGAGAAATCTTC 684  
Db 664 AAAGCTCAGCTTAATGCAACCTCAATTAATCAACAAAGATTTGGGAATGCCATCAAC 723  
QY 685 AAGAGCTTGAAGAGAGATGATGACAAATTCCTTGCATTTTGAAGTCAACCAT 744  
Db 724 AAGGATCT---GAGGCTGATCCAAATGATGAATTTGAAACCTGAGATCAAGAT 780  
QY 745 CAGTCTTG 753  
Db 781 AAGTCTTG 789

RESULT 3  
US-09-325-932A-21  
; Sequence 21, Application US/09325932A  
; Patent No. 6451604  
; GENERAL INFORMATION:  
; APPLICANT: Flinn, Barry  
; APPLICANT: Lasham, Annette  
; TITLE OF INVENTION: Compositions affecting programmed cell  
; TITLE OF INVENTION: death and their use in the modification of forestry plant develo  
; FILE REFERENCE: 1022  
; CURRENT APPLICATION NUMBER: US/09/325,932A  
; NUMBER OF SEQ ID NOS: 206  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 21  
; LENGTH: 704  
; TYPE: DNA  
; ORGANISM: Pinus radiata  
US-09-325-932A-21

Query Match 16.9%; Score 208.4; DB 3; Length 704;  
Best Local Similarity 60.4%; Pred. No. 5,1e-51;  
Matches 344; Conservative 0; Mismatches 226; Indels 0; Gaps 0;

QY 6 CGTCCGAACACTAAAGTGAAGAAATGCGACTTAAAGTTTCTGATTCGTTCC 65

Db 120 CGTGCAGATATGATTTGTCGAGAAGATGTCAACATTATATGTCGACGTTCCAAATACC 179  
QY 66 TGTCTCTTTCGATGATCTGAGCAATTGGAACCGCTTTTGAAGATGGGTATGAAACGA 125  
Db 180 GACCCCATCTGAACACTCTGAACCCCTGAGGAAGCTTTTGAAGGGGGCACAATAATGA 239  
QY 126 GCACTTATATATCATCTTTGGCTCAAGAAAGTGTGAACAGAGGAAGCATCAGACA 185  
Db 240 GAATCATATCATCAAAATATTAGACATATGAACTGACACACACGCAAAATTAATCCGTCA 299  
QY 186 AGCATACACGAAACCTACGCGCAAGACCTTCTCAAGACTCTTGACAAAGAGCTCTCTAA 245  
Db 300 AAGTTATTTTCAACTGATACGAAGAGATCTCTTGAACGATTAATAATCTGAGCTTTCAAG 359  
QY 246 CGATTTTGAGAGAGCTATCTTTGTTGTGACTCTTGAACCGGTGAGGCTGATGCTTTATT 305  
Db 360 TGACTTTGAAAGAGCTGTATCTCTTTGGGTATGATATCCAGCTGAAAGCTGAGGCTCAT 419  
QY 306 GGCCTAATGAAGCTCAAAAAGATGAGCTTGAAGCAACCAAGTTCTTATGGAAGTTGCTTG 365  
Db 420 ATCTCATGTGCATTAAGAAAGTGAATGCAAGAAATATATCGCTTTTGAATAATTCAG 479  
QY 366 CACAAGACATCAACGACGTCTTCAAGCTAGGCAAGCTTACCATGCTGCTACAAAGAA 425  
Db 480 TGTCTGATCTTGGCTGAATGATGATGATGAGGCAAGCATATCATATTCGTTACAAA 539  
QY 426 GTCTCTTGAAGAGAGCTTGTCTACCACTACCGGTGACTTCAAAAAGCTTTTGGTTTC 485  
Db 540 GTCCCTCGAAGAAAGCTGTGCTGACATACAAAGTGAAGAACTTCGTTAAGTTGCTGATGC 599  
QY 486 TCTTGTACCTCATACAGATGACGATGCAAGAGATGAAGTAACATATGCTGCTAAGCAGA 545  
Db 600 ACTTGTATGTTCAATATGATGATGAGGTCCGAAAGTGAATGATGATTTGGCAAGTTATGA 659  
QY 546 AGCTAAGCTGTGCATGAGAAATCAAGAA 575  
Db 660 AGCAAAAGAGCTAAGTAATCTATTAACGA 689

RESULT 4  
US-09-325-932A-18  
; Sequence 18, Application US/09325932A  
; Patent No. 6451604  
; GENERAL INFORMATION:  
; APPLICANT: Film, Barry  
; APPLICANT: Iasham, Annette  
; TITLE OF INVENTION: Compositions affecting programmed cell  
; FILE REFERENCE: 1022  
; CURRENT APPLICATION NUMBER: US/09/325,932A  
; NUMBER OF SEQ ID NOS: 206  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 18  
; LENGTH: 484  
; TYPE: DNA  
; ORGANISM: Pinus radiata  
US-09-325-932A-18

Query Match 15.1%; Score 186.2; DB 3; Length 484;  
Best Local Similarity 74.9%; Pred. No. 1.4e-44;  
Matches 233; Conservative 0; Mismatches 78; Indels 0; Gaps 0;  
QY 254 AGAGACTATCTTGTGTGACTCTTGAAACCGGTGAGCGTATGCTTATTTGGCTAATG 313  
Db 174 AGAGCTGTGTGCTTGTGTGCTGTATCGGTGAGGTGATGCTTATTTGGCGAATG 233  
QY 314 AACTTCAAAAAGATGACTTTCAGCAACCAAGTCTTATGGAAGTTGCTTGCACAAAGA 373  
Db 234 AAGCGAGAAAGATGACTTTCAGCAACCAAGTCTTATGGAAGTCTTGCACAGAGGT 293  
QY 374 CATCAACGAGCTTTCAGCTAGGCAAGCTTACATGCTGCTCAAGAAAGTCTCTTG 433

Db 294 CTCACAGACGTTGCTCATGTCGAAGAACAAGCATATGCTGTGATACAAAGATCGCTGG 353  
QY 434 AAGAGAGCTTGTCTACACACATACCGGTGACTTCAAAAAGCTTTTGGTTTCTTGTTA 493  
Db 354 AAGAGAGCTGCTGTACACACATGAGATTTTGTGATGTTGCTGTACCTTGTGA 413  
QY 494 CTTCAATACAGTTCGAAGAGATGAAGTGAACATGACATTTGGCTTGAAGCAAGAGCTAAG 553  
Db 414 GCTCTTACATTAATGATGAGATGAGTGAATATGACTTTGGCAAAAAGAGGCTTAAGA 473  
QY 554 TGTTCATGAG 564  
Db 474 TACTCCAGAG 484

RESULT 5  
US-09-949-016-2081  
; Sequence 2081, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2081  
; LENGTH: 1332  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-2081

Query Match 12.1%; Score 149.2; DB 4; Length 1332;  
Best Local Similarity 49.0%; Pred. No. 2.1e-33;  
Matches 456; Conservative 0; Mismatches 468; Indels 6; Gaps 2;  
QY 54 TGATTCGTCTCTGCTCTCTTGTGATGCTGAGCAATTGAGAACCGCTTTGAGAGATG 113  
Db 86 TTATTCAGACTTATAGCCCATGATGATGCTGATGAGCTATTTCAAGAAACATGAGAGAT 145  
QY 114 GGGTACAAAGAGACTTATCATATCTTGGCTCACAGAAAGTGTGAACAGAGAA 173  
Db 146 TGGAACTGATGAGAAATCTCATGCAATCTGACTGAGAGGTCAAAATGCAAGCGGCA 205  
QY 174 AGTCATCAGGCAAGCATACAGAAACCTACGCGCAAGCTTCTCAAGACTCTTGACAA 233  
Db 206 GCTGATTTGTAAGAAATATCAAGCAATGAGAAAGAGCTGAAAGATGACTTGAAGG 265  
QY 234 GGAGCTCTTAAAGATTTGAGAGAGCTATCTTGTGTGAGACTTGTGAACCGGTGAGG 293  
Db 266 TGATCTCTGTGGCACTTTGAGAGATCTATGATGAGGCTTATGATCCACAGAGCTTT 325  
QY 294 TGATGCTTATTTGGCTAATGAAAGCTTCAAAAAGATGAGACTTTCAGCAACCAAGTTTAT 353  
Db 326 TGAATGAAAGAGCTTAAAGAAATCCATGAAAGGCGCGGGAACAAAGAGATGCCCTTGA 385  
QY 354 GGAAGTGTCTTGCAAGAGCATCAACGAGGCTTTCACGCTAGGCAAGCTTACCATG 413  
Db 386 TGAATCTTAACTTACAGAGCAAGAGCAAAATGAAGATATCTTCAAGCTTATTAATC 445  
QY 414 TGGCTACAAAGAGCTCTTGAAGAGAGCTTGTCTACCACTACCGGTGACTTGAAGAA 473  
Db 446 AGTATACAAAGAGAGCTTGTGAGATGACATTTGTTCCGAATCATCTGTGATCCGGA 505  
QY 474 GCTTTGGTTTCTCTTGTTAACCTCATACAGGTACGAAGAGATGAAGTGAACATGACTT 533

Db 506 AGCTGTTGACCTTTGGCAGATGCGAAGAGATGAAAGCTTGAAGTGATGATGATCATCT 565  
QY 534 GCGTAAGCAAGAGCTTAAGTGTGCTCATGAGAAATCAAGAGCA---AGACTACAAATGA 590  
Db 566 GGGCAAAACAGAGAGCCAGATTTCTATTAAGCTGTGAGAACAGATGGGCGCGATGA 625  
QY 591 TGAGATGTTATTAAGATCTTGTCCACAAAGAACAAAGCTCAGATCAATGCTACTTTTA 650  
Db 626 AGCAAAATTCATCGAATCTCTGTGTTAAAGAGCTTTCTCAATTAATAACTAATTTGA 685  
QY 651 CCGTTACCAAGATGATCAGCGAGAAATTCGAAGAGCTTTGAGAAAGAGATGATGA 710  
Db 686 TGAATACAGAAATATCAGCCAAAGACATTTGACAGCAT---AAAAGAAATTAATC 742  
QY 711 TGCAAGTTCCTTGACATTTTGAAGTCAACCATTCAGTGTTCACAAAGCAAGAGCTTA 770  
Db 743 TGGGCAATTTTGAAGACTTACTGTGGCCATAGTTAATGTGTAGAGAACACGCGGCTT 802  
QY 771 CTGTGCGATGTTCTTGTTCAGCAATCAACAAACCTGAACTGATGAGAGACCTCAC 830  
Db 803 TTTAGCCGAAAGCTGATGAGCTTGAAGGTAATTTGAACTGATGATTTACTTTGA 862  
QY 831 TAGAATTTGACCAAGAGCTGAGATTGACTTGAAGTCAATTTGAGAGAGATGACGCG 890  
Db 863 CCGAATTAATGTGTCCAGATCAGAAATTTGACCTTTTGACATTCGAACAGAGTTCAAGAA 922  
QY 891 CAGGAACAGATTCCTTTGAGAAAGCTATTACCAAGACACTCGTGGAGATTACAGAA 950  
Db 923 GCATTTATGCTATTTCCCTATATTTCAGCAATTAATCGAATTAATCTTGAGACTATGAAT 982  
QY 951 GATGCTGTCGCACTTCTCGTGAAGATGA 980  
Db 983 CACACTTTAAATCTGTGTGAGATGA 1012

RESULT 6  
US-09-023-655-1149  
Sequence 1149, Application US/09023655  
Patent No. 6607879  
GENERAL INFORMATION:  
APPLICANT: Cocks, Benjamin G.  
APPLICANT: Susan G. Stuart  
APPLICANT: Jeffrey J. Seilhamer  
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE  
TITLE OF INVENTION: EXPRESSION  
NUMBER OF SEQUENCES: 1508  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
STREET: 3174 PORTER DRIVE  
CITY: PALO ALTO  
STATE: CALIFORNIA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/023,655  
FILING DATE: HEREWITH  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Zeller, Karen J.  
REGISTRATION NUMBER: 37,071  
REFERENCE/DOCKET NUMBER: PA-0001 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (650) 855-0555

TELEFAX: (650) 845-4166  
INFORMATION FOR SEQ ID NO: 1149:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1339 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: GENBANK  
CLONE: 9186967  
US-09-023-655-1149  
Query Match 12.1%; Score 149.2; DB 4; Length 1339;  
Best Local Similarity 49.0%; Pred. No. 2.1e-33;  
Matches 456; Conservative 0; Mismatches 468; Indels 6; Gaps 2;  
QY 54 TGATTTCTGTTCTGCTCTCTTGTGATGATGCTGAGCAATTTGAGAACCGCTTTGAAAGATG 113  
Db 88 TTATCCAGACTTTAGGCCATCAGTGATGCTGAAGCTATTCAAGAAAGCAATCAGAGGAAT 147  
QY 114 GGGTACGAAGAGGACTTGATGATATCAATCTTGGTCAAGAGTGTGAACAGAGAA 173  
Db 148 TGAATGATGAGAAATGCTCATCAGCATTTCTGACTGAGAGTCAATGACACAGCGCA 207  
QY 174 AGTCATCAGCAAGCATACCAAGAAACCTACGCGAAGACCTTTCAGACCTTTGACAA 233  
Db 208 GCTGATTTTGAAGATATCAAGCAGCATATGAAAGAGCTGAAGATGACTTGAAGGG 267  
QY 234 GGAAGCTCTTAAGATTTGAGAGACTATCTTTGTGTGAGACTTTGAAACCGGTGAGGG 293  
Db 268 TGATCTCTGCGCACTTTGAGCATCTCATGTGTGCCCTTAGTACTCCACGACGCTT 327  
QY 294 TGATGCTTATTTGGCTAATGAGCTACAAAGAAAGATGAGACTTCAAGCAACCAAGTCTTAT 353  
Db 328 TGATGCAAGCAGCTTAAGAAATCCATGAGGCGCGGGAACAAAGAGATGCTTGAAT 387  
QY 354 GGAAGTGTCTTCAAGAGCATCAACGAGCTGCTTCAAGCTAGCAAGCTTACCATGC 413  
Db 388 TGAAATCTTAATCAACAGCAAGCAAGCAATGAAGATATCTCTCAAGCTTATATATC 447  
QY 414 TGGCTTCAAGAGTCTTTGAAGAGAGCTTGTCTACCACTACCGGTGACTTTCAGAA 473  
Db 448 AGTATTAAGAAAGAGCTTTGAGATGACATTTGTTCCGAACATCTGTGACTTCGGAA 507  
QY 474 GCTTTTGGTTCTCTTGTGATCTCATACAGTACAGAGAGATGAGATGATGATCAT 533  
Db 508 AGCTCTGTGACTTTGGCAGATGCGAAGAGATGAAAGTCTGAAAGTGAATGAGATCT 567  
QY 534 GCGTAAGCAAGAGCTTAAGCTGTGCTCATGAGAAATCAAGACA---AGACTACAAATGA 590  
Db 568 GGGCAAAACAGAGAGCCAGATTTCTATTAAGCTGTGTGAGAACAGATGGGCAACGATGA 627  
QY 591 TGAGATGTTATTAAGATCTTGTCCACAAAGAACAAAGCTCAGATCAATGCTACTTTTA 650  
Db 628 AGCAAAATTCATCGAATCTCTGTGTTAAAGAGCTTTCTCAATTAATACTAATTTGA 687  
QY 651 CCGTTACCAAGATGATCAGCGAGAAATTCGAAGAGCTTTGAGAAAGAGATGATGA 710  
Db 688 TGAATACAGAAATATCAGCCAAAGACATTTGAGACAGCAT---AAAAGAAATTAATC 744  
QY 711 TGCAAGTTCCTTGACATTTTGAAGTCAACCATTCAGTGTTCACAAAGCAAGAGCTTTA 770  
Db 745 TGGGCAATTTTGAAGACTTACTGTGGCCATAGTTAATGTGTGAGAACACGCGGCTT 804  
QY 771 CTGTGCGATGTTCTTGTTCAGCAATCAACAAACCTGAACTGATGAGAGAGCATCTCAC 830  
Db 805 TTTAGCCGAAAGCTGATGAGCTTGAAGGCTTTGAACTGATGAGATTTTACTTTGA 864  
QY 831 TGAATTTGACCAAGAGCTGAGATTGACTTTGAAAGTCAATTTGAGAGAGATGACGCG 890  
Db 865 CCGAATTAATGTGTCCAGATCAGAAATTTGACCTTTTGGACATTTCCAAAGAGATTTCAAGAA 924  
QY 891 CAGGAACAGATTCCTTTGAGAAAGCTATTACAAAGACACTCGTGAAGATTACAGAA 950

Db 925 GCATTATGCTATTCCCTTATTTCAGCAATTAAATCGATTCTTGAGAGCTATGAAT 984  
 951 GATGCTCGTGCACCTTCGCTGAGATGA 980  
 Db 985 CACACTCTTAAAAATCTGTGTGAGATGA 1014

RESULT 7  
US-09-94

```

: Sequence 722: Application US/09949016
: Patent No. 6812339
: GENERAL INFORMATION:
: APPLICANT: VENTER, J. Craig et al.
: TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
: TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
: FILE REFERENCE: CLO01107
: CURRENT APPLICATION NUMBER: US/09/949,016
: CURRENT FILING DATE: 2000-04-14
: PRIOR APPLICATION NUMBER: 60/241,755
: PRIOR FILING DATE: 2000-10-20
: PRIOR APPLICATION NUMBER: 60/237,768
: PRIOR FILING DATE: 2000-10-03
: PRIOR APPLICATION NUMBER: 60/231,498
: PRIOR FILING DATE: 2000-09-08
: NUMBER OF SEQ ID NOS: 207012
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 722
: LENGTH: 1339
: TYPE: DNA
: ORGANISM: Human
: US-09-949-016-722

```

Query Match	12.1%	Score 149.2	DB 4	Length 1339
Best Local Similarity	49.0%	Pred. No. 2.1e-33		
Matches 456; Conservative	0;	Mismatches 468;	Indels 6;	Gaps 2;

Oy		54	TGATTTCTGTTCCTGCTCCTTCTGGATGATGCTGAGCAATTGGAACCGCTTTTGAAAGATG	113
Dd		88	TTATCCAGACTTTTAGCCCATCAGTGGATGCTGMAAGCTTTAGAAAGCAATCAGAGGAT	147
Oy		114	GGGTAGCAACGAGGACTTGATCATATCAATCTTGGCTCACAGAAGTGTGAACAGAGAA	173
Dd		148	TGGAACATGATGAAAAATGCTCATCAGCACTTCTGACTGAGAGGTCAAATGCAACGCCGA	207
Oy		174	AGTATCATGGCAAGCATTCACAGAACTTAAGCGGAGAAGACTTTCAGAACTCTTGACA	233
Dd		208	GCTGATTTGTTAAGGATAATCAAGACGCTAATGAAAGAGCTGAAGATGACTTGAAGGG	267
Oy		234	GGAGCTCTCTAACATTTGAGAGAGCTATCTTGTGTGGA CTCTGAAACCGGTAAGCG	293
Dd		268	TGATTTCTCTGGCCACTTTGAGCATCTCATGTGTGCGCTTAAGTACTCCACCAGCAGTTTT	327
Oy		294	TGATGCTTTATTTGSCATAAGAGCTACAAAAGATGCACTTCAAGCAACCAAGTCTTAT	353
Dd		328	TGATGCAAAAGCAGTTAAAGAAATTCATGAAGGGCGGGGAACAAACGAATGCTTGAT	387
Oy		354	GGAAAGTCTTGACCAAGAGATCAACGCACTGCTTACGCTAGGCAAGCTTACATGC	413
Dd		388	TGAAATCTTAACTTACAGGACAAACAGGCAATGAMAGATTTCTTCACAGCTTATTATC	447
Oy		414	TGCGTACAGAAGCTCTTGAAGAAGACGTTGCTCACACACTACCGSTGACTTCAGAA	473
Dd		448	AGTTATCAAGAGAGCTTTGGAGATGACATTATGTTCCGAACATCTGGTGACTTCCGGA	507
Oy		474	GCTTTTGGTTCTCTTGTTAACCTATACAGGTACGAAGAGATGAAAGTGAACATGACATT	533
Dd		508	AGCTCTGTGACTTTGGCAGATGGGAGAAAGATGAAGAAGTGAAGTGAAGATGACATCT	567
Oy		534	GGCTTAAGCAAGACTAAGCTGTTCATGAGAAAATCAAGACA--AGCATCAATGA	590
Dd		568	GGCCAAACAGAGTCCAGATTCTCTTAAAGCTGTGTGAAGACAGATGGGCAACGATGA	627

Qy	591	TGAGATGTTATTGAATCTTGTCCAGAAAGCAAAAGCTCAGATCAATGCTACTTTAA	650
Db	628	AGCAAAATTCACGTGAGATCTGTGTTTAAAGAGCTTCTCCTAAATTAACCTAACTTTGA	687
Qy	651	CCGTATCCAAAGATGATCATGTGCGAGAAATTCCAAGAGCTTGAGAGAAAGAGATGATGA	710
Db	688	TGAAATCAGAAATATCAGCCAAAAGACATTTGTGACAGCAT---AAAGAGAAATATC	744
Qy	711	TGACAAGTTCCTTGACATTTGAGGTCAACCATTCAGTGTCTGACAAAGACCAAGCTTTA	770
Db	745	TGGGCAATTTGAAAGCACTTACTGTTGGCATATGTTATGTGTGAGGAAACGCGGCTT	804
Qy	771	CTTTGCGATGTTCTTGCTGTGAGCAATCAAAACTGGAATGATGAAGAGACATCAC	830
Db	805	TTTAGCGCAAGACGTGCATCGAGCTTTAAGGGTATGGAATGATGATGTTACTTGAA	864
Qy	831	TAGAAATGTGACCAAGAAGCTGAGATGACTTGGAAGTCAATTGAGAGGAGTACCAAG	890
Db	865	CGAAATTAATGTGTCCAGATCAGAAATTAAGACCTTTTGGACATTTGAAACAGATTCAGAA	924
Qy	891	CAGAAACAGCATTCCTTTGAGAAAGCTATTAACCAAGACATCTGTGAGATTACGAA	956
Db	925	GCATTTGTGCTATTCCTATATTCAGCAATTAATCGGATCTTCTGAGACTATGAAT	984
Qy	951	GATGCTGTCGACATTCCTCGTGAAGATGA	980
Db	985	CACACTTTAAAAATCTGTGTGAGATGA	1014

## RESULT 8

```

US-09-325-932A-19
: Sequence 19, Application US/09325932A
: Patent No. 6451604
: GENERAL INFORMATION:
: APPLICANT: Flinn, Barry
: APPLICANT: Isaham, Annette
: TITLE OF INVENTION: Compositions affecting programmed cell
: TITLE OF INVENTION: death and their use in the modification of forestry plant devel
: FILE REFERENCE: 1022
: CURRENT APPLICATION NUMBER: US/09/325,932A
: CURRENT FILING DATE: 1999-06-04
: NUMBER OF SEQ ID NOS: 206
: SOFTWARE: FaastSeq for Windows Version 3.0
: SEQ ID NO 19
: LENGTH: 221
: TYPE: DNA
: ORGANISM: Pinus radiata
US-09-325-932A-19

```

Query Match	11.3%;	Score 138.4;	DB 3;	Length 221;
Best Local Similarity	76.8%;	Pred. No. 1.1e-30;		
Matches 169;	Conservative	0;	Mismatches 51;	Indels 0;
			Gaps	0

QY	299	CTTTATGGCTAATGAAGCTACAAAAGATGAGCTTACAGCAACCAAGTTCTTATGGAAG	358
Db	1	CGTACTGGCGAATAGACGAGAAAGATGAGACTTACAGCAACAGGTTCTTATGAAAA	60
QY	359	TTGCTTGCACAAGGACATCAACGACGCTGTCAACGCTAGGCAAGCTTACATGCTCGCT	418
Db	61	TAGCTCGACGAGGCTCTCCGACGAGTTGCTTATGGAAGCAAGCAATCATGCCGAT	120
QY	419	ACAAGAAGCTCTTAAGAGGACGTGTCAACAACGTACCGGGAATTGAGAAAGCTTT	478
Db	121	ACAAGAAGTCGCTGAAGAAGACGTGGCTACCAACAACCTGAGATTTGCTAAGTTGC	180
QY	479	TGGTTTCTCTTGTACCTCATACAGGTACGAAGAAGATGA	518
Db	181	TGGTACCTCTTGTAGCTCTCAACGTTATGATGATGAGATGA	220

RESULT 9  
US-09-023-655-1152  
; Sequence 1152, Application US/09023655

```

; Patent No. 6607879
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Susan G. Stuart
; APPLICANT: Jeffrey J. Sellhammer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
; TITLE OF INVENTION: EXPRESSION
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/023,655
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0001 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1152:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1362 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: g187137
; US-09-023-655-1152

```

```

Query Match      8.4%; Score 103.8; DB 4; Length 1362;
Best Local Similarity 46.7%; Pred. No. 5.3e-20;
Matches 435; Conservative 0; Mismatches 487; Indels 9; Gaps 3;

```

```

QY 56 ATTCTGTCCTGCTCTCTGATGATGCGAATTTGAACCGCTTTTGAAGATGG 115
DB 138 ATACTAATTGATGCTGAGCGGATGCTTTGAACATTGAACAGCATCAACCAAG 197
QY 116 GTAAGACGAGGACTTATCATATCATCTTGGCTCAACAAGTGTGAACAGAGAAAG 175
DB 198 GTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 257
QY 176 TCATCAGGAGAGATACCAAGAACTTACGCGAAGACCTTCTCAAGACTTTGAACAG 235
DB 258 ATATTGCTCTGCTCTGCTCTGATGATGCGAATTTGAACCGCTTTTGAAGATGG 317
QY 236 AGCTCTCTAAGATTTGAGAGAGCTATCTGTTGGAATCTTGAACCGGAGAGCG 295
DB 318 CCTTATCTGCGCCTGAGAGCGTATTTTGGGCTTATGAAGACATCTGCTCAATG 377
QY 296 ATGCTTATTTGCTTATGAGCTACAAAGATGAGCTTCAAGCAACCAAGTTCTTATG 355
DB 378 AGCTTCTGAGCTTAAAGCTTCAATGAGGCGCTGGGAACCGAGGAGCTCTCATG 437
QY 356 AAGTTCCTTGCACAGAGCATCAACGAGCTGCTTCAACGCTAGGAGCTTACCATG 415
DB 438 AGATCATCTGCTCTCAGAACCAACGAGAGCTGAGGAAATTAACAGATCTACAGAA 497

```

```

QY 416 GCTACAGAAGTCTCTTGAAGAGAGCGTTGCTCACCACTACCGGTGACTTCAGAAAGC 475
DB 498 TGTACAGAAGTCTCTTGAAGAGAGCGTTGCTCACCACTACCGGTGACTTCAGAAAGC 557
QY 476 TTTGTTTCTCTTGTATCTTATCTTATCTTATCTTATCTTATCTTATCTTATCTTAT 532
DB 558 TGATGTTGCTCTGCGAAAGGTTAGAAAGCAGAGATGCTCTGCTCATTTATTTGAAC 617
QY 533 TGGCTAAGAGAGAGCTTAACTGCTGCTCATGA--GAAATCAAGACCAAGCACTACATG 589
DB 618 TGATTGACCAAGATCTGCGATCTTATGAGCTGCTGAGTGAAGGAAAGAACTGATG 677
QY 590 ATGAGATCTTATTTGAATCTTGTCCACAAGAGCAAGCAAGCTGATTCATCTTATTTA 649
DB 678 TTCCAGAGTGAATGACATCATGACCGAGCGAGCGTCCCACTCCAGAAAGATTTTG 737
QY 650 ACCGTTACCAAGATGATCATGCGAGAAATTTCTAAGATCTTGAAGAGAGATGATG 709
DB 738 ATAGGTACAAAGATTACAGCCCTTATGACATGTTGAAAGCATCAGAAAGAGTTAAAG 797
QY 710 ATGACAAGTCTTGTGATCTTGTGATGATGATGATGATGATGATGATGATGATGAT 769
DB 798 GAGACCTGGAAGATGCT--TTTCTGAACCTGTTCAAGTCAATGACAAAGCCCTGT 854
QY 770 ACTTGTGATGCTTCTGCTTCAAGCAATCAACAAATGSAATGATGAGAGCACTCA 829
DB 855 ATTTGCTGATGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 914
QY 830 CTAGAATTTGACCAACAAGAGCTGATGATGATGATGATGATGATGATGATGATGAT 889
DB 915 TCAGATCATGCTCTCCGCAAGTGAAGTGAATGATGATGATGATGATGATGATGATGAT 974
QY 890 GAGAGAAAGCATCTTCTTTGAGAAAGCATATTCACAAAGCATCTGAGATTTCCAGA 949
DB 975 GAAAGTACGCGAAGTCTCCGTAATCTTATTCACAGCAAGCATGAAAGCGCATCCAGA 1034
QY 950 AGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 980
DB 1035 AAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1065

```

```

RESULT 10
US-09-814-915A-101
; Sequence 101, Application US/09814915A
; Patent No. 6750015
; GENERAL INFORMATION:
; APPLICANT: Horwitz, Kathryn
; APPLICANT: Richer, Jennifer
; TITLE OF INVENTION: Progesterone Receptor-Regulated Gene Expression and Methods Relat
; FILE REFERENCE: 2848-39
; CURRENT APPLICATION NUMBER: US/09/814,915A
; PRIOR FILING DATE: 2002-03-21
; PRIOR APPLICATION NUMBER: 60/214,870
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 101
; LENGTH: 1362
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-814-915A-101

```

```

Query Match      8.4%; Score 103.8; DB 4; Length 1362;
Best Local Similarity 46.7%; Pred. No. 5.3e-20;
Matches 435; Conservative 0; Mismatches 487; Indels 9; Gaps 3;

```

```

QY 56 ATTCTGTCCTGCTCTCTGATGATGCGAATTTGAACCGCTTTTGAAGATGG 115
DB 138 ATACTAATTGATGCTGAGCGGATGCTTTGAACATTGAACAGCATCAACCAAG 197
QY 116 GTAAGACGAGGACTTATCATATCATCTTGGCTCAACAAGTGTGAACAGAGAAAG 175

```

Db 198 GTGTGATGAGGTCAACATTTGTCACATTTTGAACCAACCGCAGCAATGCACAGAGCAGG 257  
 Oy 176 TCATCAGGCAAGCATACACGAAACCTACCGCGAGAACCTTCTCAAGACTCTTGAACAAG 235  
 Db 258 ATATTGCTTCGCTACACAGAGAGGACCAAAAAGAACCTTGATCAGCATGAAGTCAAG 317  
 Oy 226 AGCTCTTACGATTTTCAGAGAGCTATCTTGTGTGAGACTCTTGAACCCGGTAGCGTG 295  
 Db 318 CCTTATCTGGCCACCTCGAGAGACGGTGATTTTGGGCTTATTTGAAGACACCTGCTCAGTATG 377  
 Oy 226 ATGCTTTATTTGGCTATGAGGCTAACAAAAGATGACCTTCAAGCAACCAAGTCTTATGG 355  
 Db 378 AGCCTTTCGAGCTAAAGCTTTCATGAGAGGGGCTGGGAACCGACGAGACTCTTCAATG 437  
 Oy 356 AAGTTGCTTGCAACAAGGACATCAACGACAGCTGCTTCAACGCTAGGCAAGCTTACATGCTC 415  
 Db 438 AGATCATCTGCTCCAGAACCAACGAGAGCTGAGGAAATTTAACAAGTCTTACAAGGAAA 497  
 Oy 416 GCTACAGAAGATCTCTTGAAGAGAGAGCTTGCTCAACAACCTACCGGTGACTTGAAGAGC 475  
 Db 498 TGTACAAGACTGATCTGAGAGAGACATTATTTCCGACACATCTGGTGTGACTTCCGAAAG 557  
 Oy 476 TTTTGGTTTCTGTTTACTCATACAG---GTACGAGAGAGATGAAAGTAATGACAT 532  
 Db 558 TGATGCTGCTCCCTGGCAAGGGGTAGAGAGAGAGAGATGCTCTGTGATTGATTAAGAAC 617  
 Oy 533 TGGCTAAGCAAGAGCTAAGCTGCTGATCATA---GAAATATCAAGAGCAAGACATCAATG 589  
 Db 618 TGATTTGACCAAGATGCTCGGAGATCTTATGACGCTGAGATGAGAGAGAGAAAGAACTGATG 677  
 Oy 550 ATGAGAGATGTTATTGAATCTTTGTCCACAAGAGCAAGAACTCAGATCAATGCTACTTTTA 649  
 Db 678 TTCCCAAGTGATCAGCATCATGACCGAGCGAGAGCTGCCCACTTCCAGAAAGTATTTG 737  
 Oy 650 ACCGTTACCAAGATGATCATGCGAGAGAAATTTCTAAGAGTCTTGAGGAAGAGATGATG 709  
 Db 738 ATAGGTACCAAGATTACAGCCCTTATGACATGTTGAAAGCAACAGAAAGAGTTAAAG 797  
 Oy 710 ATGACAAGTTCCTTGACATTTTGAAGGTCAACACATTCAGTCTTGACAAAGCCAGAGCTTT 769  
 Db 798 GAGACCTGGAATAATGCT--TTTCTGTAACCTGTTTCAAGTGCATTTCAAGAACAGCCCTGT 854  
 Oy 770 ACTTGTGCTGATGCTTCTTCTTGACGAATCAACAAAACCTGAACTGATGAGAGACACTCA 829  
 Db 855 ATTTTGTGATGCGGCTGATGACTCATGAGAGGCAAGGGAGCGGAGATTAAGGTCTGA 914  
 Oy 830 CTAGAATTGTGACACACAGAGCTGAGATTGACTTGAAGGTCAATGAGAGAGATACAGC 889  
 Db 915 TCAGATCATGTGCTTCCCGCAGTGAAGTGAACATGTTGAAATTTAGGTCTGAATTCAGAG 974  
 Oy 890 GCAGGAACAAGCACTTCCTTTGGAGAAAGCTATTACAAAAGACACTCGTGAAGATTAGAG 949  
 Db 975 GAAAGTTAGCGCAAGTCCCTGTAATTAATTAATTCAGCAAGACATAAGGGGACTACACAG 1034  
 Oy 950 AGATGCTGCTGCGCACTTCTCGGTGAAGATGA 980  
 Db 1035 AAGCGCTGCTGTAACCTGTGTGTTGGTGAAGATGA 1065  
 RESULT 11  
 US-09-949-016-2431  
 ; Sequence 2431, Application US/0949016  
 ; Patent No. 6812339  
 ; GENERAL INFORMATION:  
 ; APPLICANT: VENTER, J. Craig et al.  
 ; TITLE OF INVENTION: POLYMOBPHISMS IN KNOWN GENES ASSOCIATED  
 ; FILE REFERENCE: CL001307  
 ; CURRENT APPLICATION NUMBER: US/09/949,016  
 ; CURRENT FILING DATE: 2000-04-14  
 ; PRIOR APPLICATION NUMBER: 60/241,755  
 ; PRIOR FILING DATE: 2000-10-20

	Query Match	8.4%	Score 103.8	DB 4	Length 1614
Qy	ATTCTGTTCCGCTCCTTCTGATGATGCTGAGCAATTGAGAACCGCTTTTGAAGATGGG	46.7%	Pred. 5.9e-20		
Db	205 ATTACTAATTGATGCTGAGCGGGATGCTTTGAAACATTGAAACAGCCATCAAGACCAAG	0	Mismatches 487	Indels 9	Gaps 3
Qy	116 GTACGAACGAGAGCTTGATCATATTCATCTTGGCTCAAGAGTGTCTGAACAGAGAAAG				
Db	265 GTGAGATGAGAGTCAACATTTGTCAACATTTTGACCAACCGCAGCAATTCACAGAGACAGG				
Qy	176 TCATCAGGCAAGCATCCACGAAACCTAAGCGGCAAGACCTTCTCAAGACTTTGACAAAG				
Db	325 ATATTGCTCTTGCCTCCACGAGAAAGAACCAAAAAGAACTTGATCAGACCTGAAGTCA				
Qy	236 AGCTCTCAAGATTTGAGAGAGACTATCTTGTGTGAGACTCTTGAACCCGCTGAGCGTG				
Db	385 CCTTATCTGACCTGAGACGGTGAATTTTGGGCTATTGAAGACACTCTCAGATATG				
Qy	296 ATGCTTTATTTGGCTAATGAACTCAAAAAGATGAGACTTTCAGAACCAACAAATTTCTATGG				
Db	445 AAGCTTCTGAGCTAAAGCTTCCATGAAGGGGCTGGAAACCGAGAGGACTCTCATTTG				
Qy	356 AAGTTGCTTGCAAGAGCATCAACGCACTGCTTCACTGCTAAGGCAAGCTTACATGCTC				
Db	505 AGATCATCTGCTCCAGAACCAACGAGGAGCTGCGAGAAATTTACAGATCTTACAGAGAA				
Qy	416 GCTCAAGAAATCTCTTGAAGAGACGTTGTCAACACATACCGGTACTTTCAGAAAGC				
Db	565 TGTCAAAAGCTGATCTGAGAAAGACATTAATTTGGAAACATCTGTGTACTTCCGCAAGC				
Qy	476 TTTTGGTTTCTCTTGTACTCATACAG---GTAAGAAAGATGAAATGACAT				
Db	625 TGAATGTTGCTTCCGCAAGAGGTGAAAGAGAGAGATGCTCTGTCAATTGATTAAGAAC				
Qy	533 TGGCTAAGCAAGAGCTAAGCTGTGCTCAATG--GAAATCAAGGACCAAGCATCAATAG				
Db	685 TGATTTGACCAAGATGCTGGGATCTTATATGAGCTGTGAGTAAAGAGAAAGAACTGATG				
Qy	590 ATGAGATGTTATTAGATCTTTGTCCACAAAGCAAGCTCAATCAATGCTACTTTTA				
Db	745 TTCCCAATGATCAGATCATGACCGAGCGAGCGTGCCTCCACTTCAGAAAGTATTTG				
Qy	650 ACCGTTACCAAGATGATCATGCGCAGGAAATTTCAAGATCTTGAAGAGAGATGATG				
Db	805 ATAGATACAAAGATTACAGCCCTTATGACATGTTGAAAGCATCAGGAAAGAGTTAAG				
Qy	710 ATGACAAAGTCTTGACATTTTGAAGTCAACCATTCAGTGTCTTGACAAAGACAGACTTT				
Db	865 GAGACCTGGAATAATG---TTTCTGAAACCTGGTTCAATGATTCAGAAACAAGCCCCGT				
Qy	770 ACTTTGTGATGTTCTTCTGTTCAAGCATTCACAAAACTGGAACGTATGAAAGAGACATCA				
Db	922 ATTTTGTGATCGGCTGTATGATCTCAATGAAGGCAAGGGGACCGGATTAAGTCTCGA				
Qy	830 CTAGAATTGTGACCAAGAGCTGAGATTGACTTGAAGGTATGTTGAGAGGATGATACAGC				
Db	982 TCAAGATCATGCTCTCCGCAAGTAAAGTGAATGATGTTGAAATTAAGTCTGAAATTCAGA				
Qy	890 GCAGGAACGACATTTCTTTGGAGAAAGCTATTACAAAGACATCTGTGAGATTACGAGA				



```
Db      1042 GAAAGTACGGCAAGTCCCTTACTATTATATCCAGCAAGACACTAAAGGCGACTACAG 1101
Qy      950 AGATGCTGTCGCACTTCTGCGTGAAGATGA 980
Db      1102 AAGCGTCTGCTACCTGTGTGTGAGATGA 1132

RESULT 12
US-09-919-039-393
; Sequence 393, Application US/09919039
; Patent No. 6727066
; GENERAL INFORMATION:
; APPLICANT: Kaseer, Matthew R.
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED HUMAN C3A LIVER CELL CULTURES
; FILE REFERENCE: PP-0035 US
; CURRENT APPLICATION NUMBER: US/09/919,039
; PRIOR FILING DATE: 2002-09-09
; PRIOR APPLICATION NUMBER: 60/222,113
; NUMBER OF SEQ ID NOS: 401
; SOFTWARE: PERL Program
; SEQ ID NO 393
; LENGTH: 1660
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6727066 1326902.6
US-09-919-039-393

Query Match      8.4%; Score 102.8; DB 4; Length 1660;
Best Local Similarity 46.7%; Pred. No. 1.2e-19;
Matches 435; Conservative 0; Mismatches 487; Indels 10; Gaps 3;

Qy      56 ATTCTGTTCTCTCTCTCTCTGATGATGCTGAGCAATTGAGAACCCGTTTGAAGATGG 115
Db      238 ATACTAATTGATGCTGAGCGGGATGCTTTGAACATTGMAACAGCATCAACCAAG 297
Qy      116 GTACGAACGAGGACTTGCATCATCATCTTGGCTCAAGAACTGCTGAACAGAGAAAG 175
Db      298 GTGTGATGAGTGCACCATTTGCAACATTTTGACCAACGCAACAAATGACAGAGACAG 357
Qy      176 TCATCAGGCAAGATACCAAGAACTACGCGCAAGACCTTTCAGAGCTCTTGACAAAG 235
Db      358 ATATTGCTTCGCTGCTACAGAGAGAGACCAAAAGAACTGTGATGACAGCTGAATCAG 417
Qy      236 AGCTCTCTTAAGATTTCGAGAGAGCTATCTTTGTGGACTTTGAACCCGGTGAACG 295
Db      418 CCTTATCTGGCCACCTGAGAGACGTTGATTTTGGGCTTATGGAAGACACCTGCTCAG 477
Qy      296 ATGCTTTATTTGGCTAATGAGCTACAAAAGATGACTTCAAGCAACCAAGTTCTTA 355
Db      478 ACGCTTCTGAGTAAAGCTTCCATGAAGGGGCTGGAAACGACGAGGACTCTCTCAT 537
Qy      356 AAGTCTCTGACAAAGACATCAACGACAGCTGCTTACGCTAGGCAAGCTTACCATGTC 415
Db      538 AGATCATCTGCTCCAGAACCAACAGAGAGTGCAGAAATTAACAAGTCTACAGAGAA 597
Qy      416 GCTACAAAGAGTCTCTTGAAGAGAGCTTGTCAACAACCTACCGTGAACCTTCAAAAG 475
Db      598 TGTACAAAGACTGATCTGAGAAAGACATTATTTCCGACACATCTGTGACTTCCGAC 657
Qy      476 TTTTGTCTCTCTTTTACTCTCATACAG---TAAAGAGAGATGAATGAATGAATGA 531
Db      658 TGAATGTGCGCTTGGCAAAAGGTGAAGAGAGAGATGGCTCTGTCAATTGATTGA 717
Qy      532 TTGGCTAAGCAAGAGTGAAGTGTGCATGA---GAAATCAAGGACAAGCATCAAT 588
Db      718 CTGATTGACCAAGATGCTGGGATCTCTATGACGCTGAGTGAAGGAAAGAACTGAT 777
Qy      589 GATGAGAGATTATTAGAACTTTGTCCACAAGAGCAAAAGCTCAGATCAATGCTACTTT 648
```

```
Db      778 GTTCCCAAGTGATCAGATCATGACCGAGCGGACGTGCCCACTCCAGAAATATTT 837
Qy      649 AACCGTTACCAAGATGATCATGACGAGAAATTTCAAGATCTTGAGAAAGAGATGAT 708
Db      838 GATTAAGTACAAAGATTAACAGCCCTTATGACATGTTGAAAGATCAGAAAGAGTTAA 897
Qy      709 GATGCAAGTCTCTTGAACCTTTTGAAGTCAACCATTTACGTGCTTGAACAAGACGCTT 768
Db      898 GGAGACCTGGAAATATG---TTTCCTGAACCTGGTTAGTGTGATTCAGAAACAGCCCTG 954
Qy      769 TACTTTGTGATGTTCTTCTGTTCAAGCAATCAAAACCTGAACTGATGAAGAGACATC 828
Db      955 TATTTTGTGATGCGCTGTATGATCTCATGAAGGCGCAAGGCGGACCGAGATTAAGTCTG 1014
Qy      829 ACTGAATTTGATGACCAAGAGCTGAGATGATCTTGAAGAGTCAATTTGAGAGATGACAG 888
Db      1015 ATCAGATATGATGTTCCCGAGTGAAGTGAAGATGTAATTAATGATCTGAATTTCAAG 1074
Qy      889 CCGAGAACAGCATTCCTTTGAGAAAGCTATTACCAAGACACTGCTGAGATTAACGAG 948
Db      1075 ABAAGTACGCAAGATCCCTGTACTATATATCCAGCAAGACACTAAGGGGACTACAG 1134
Qy      949 AAGATGCTGTCGCACTTCTCGTGAAGATGA 980
Db      1135 AAGCGCTGCTGTACTGTGTGTGAGATGA 1166
```

```
RESULT 13
US-09-324-096A-1
; Sequence 1, Application US/09324096A
; Patent No. 6323313
; GENERAL INFORMATION:
; APPLICANT: Tait, Jonathan
; APPLICANT: Brown, David
; TITLE OF INVENTION: ANNEXIN DERIVATIVE WITH ENDOGENOUS CHELATION SITES
; FILE REFERENCE: USFW-1-13841
; CURRENT APPLICATION NUMBER: US/09/324,096A
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patent version 3.0
; SEQ ID NO 1
; LENGTH: 981
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(981)
US-09-324-096A-1
```

```
Query Match      8.3%; Score 101.6; DB 3; Length 981;
Best Local Similarity 46.0%; Pred. No. 1.9e-19;
Matches 418; Conservative 0; Mismatches 484; Indels 6; Gaps 2;

Qy      76 GATGATGCTGAGCAATTGAGAACCGCTTTTGAAGATGGGTACGAAAGAGACTTGA 135
Db      76 GCTGATGAGAAACTCTTCGGAAGGCTATGAAGGCTTGGCAAGATGAGAGAGATC 135
Qy      136 ATATCAATCTTGCTCAAGAAAGTCTGAAGAGAGAAAGTCAATCAGGCAAGCATACAC 195
Db      136 CTGACTCTGTGACATCCGAAAGTATGCTCAGGCGCCGAGAAATCTCTGACGCTTTAA 195
Qy      196 GAAACCTACGCGGAAAGACCTTCTCAAGACTCTTGAAGAGAGAGCTCTAAGATTGAG 255
Db      196 ACTCTGTTTGGCAGGAGATCTTCTGATTAACCTGAATACAGAACTACTGAGAAATTTGA 255
Qy      256 AGAGCTATCTTGTGTGACTTTGAACCCGATGAGCTGTATTTGAGCTTAATGA 315
Db      256 AATTAATTTGAGCTCTGATGAAGCCCTGTGGCTTTATGATGCTTATGAATGAACAT 315
Qy      316 GCTACAAAGAGTGAAGCTTCAAGCAACCAAGTTCTTATGAAGTTGCTTGCAAGAGACA 375
Db      316 GCTTGAAGGAGGCTGGAACAAATGAAGAAATGACTGACAGAAATATATGCTTCAAGACA 375
```



;/ CURRENT FILING DATE: 1999-06-01  
;/ NUMBER OF SEQ ID NOS: 12  
;/ SOFTWARE: PatentIn version 3.0  
;/ SEQ ID NO 5  
;/ LENGTH: 981  
;/ TYPE: DNA  
;/ ORGANISM: Homo sapiens  
;/ FEATURE:  
;/ NAME/KEY: CDS  
;/ LOCATION: (1)..(981)  
US-09-324-096A-5

Query Match 8.3%; Score 101.6; DB 3; Length 981;  
Best Local Similarity 46.0%; Pred. No. 1.9e-19;  
Matches 418; Conservative 0; Mismatches 484; Indels 6; Gaps 2;

QY 76 GATGATGCTGAGCAATTGAGACCGCTTTGAAAGATGGGGTACGAAACGAGACTTGATC 135  
DB 76 GCTGATGCAAGAACTCTTCGGAAGGCTATGAAAGGCTTGGGCAAGATGAGAGACATC 135  
QY 136 ATATCAATCTTGGCTCAGACAGAGTGTGAACAGAGAAAGTCATCAGGCAAGCATACAC 195  
DB 136 CTGACTCTGTGACATCCGAGATTAATGCTCAGCGCCAGAAATCTCTGACGCTTTTAAG 195  
QY 196 GAAACCTACGCGCAGACCTTCTCAAGACTCTTGACAAGGAGCTCTTAACGATTTGAG 255  
DB 196 ACTCTGTTGGCAGGGATCTTCTGATGACTGAATCAGAACTAAGTGGAAATTTGAA 255  
QY 256 AGAGCTATCTTGTGTGAGACTTTGAACCCGTTGAGCGTGAATGCTTTATTTGGCTAATGA 315  
DB 256 AATTAATGTGTGCTGATGAACCTCTCGGCTTTATGATGATGAACTGAAACAT 315  
QY 316 GCTACAAAAGATGAGACTTGAAGCAACAGTCTTATGGAAGTGTGCTGACAAGAGCA 375  
DB 316 GCCTTGAGGAGCTGGAACAAATGAAGAGTACTGACGAATTAATGCTTCAAGAGCA 375  
QY 376 TCAACGAGCTGCTTCAAGCTAGGCAAGCTTACCATGCTCGTACAAGAAATCTTTGAA 435  
DB 376 CTTGAAGAACTGAGAGCATCAACAAAGTTTATGAAGAAATATGAGCTCAAGCTGAA 435  
QY 436 GAGAGCTTGTCTACCACTACCTGCTGACTTCAAGAAAGCTTTTGTCTCTTTACC 495  
DB 436 GATGACGTGTGGGGGCACTTCAGGGTACTTACAGCGATGTTGGTGTCTCTTCAAG 495  
QY 496 TCATACAGGTAGCAAGAGATGAAGTGAATGATCATGCAATGCTAAGCAAGCTAAGCTG 555  
DB 496 GCTAACAGAGACCTGTGCTGGAATTTGATGAAGCTCAAGTTGAACAAGATGCTCAGGCT 555  
QY 556 GTCATAGAGAAATCAAGACAAAGCACTACATGATGAGG--ATGTTATTAGAAATCTTG 612  
DB 556 TTATTTCAAGCTGAGAGACTTAATGGGGGACAGATGAAGAAAGTTTATCACCATCTTT 615  
QY 613 TCACAAGAAAGCAAGCTCAGATCAATGCTACTTTTAAACCGTTACCAAGATGATCATGGC 672  
DB 616 GGAACACGAAGTGTCTCATTTGAGAAAGGTGTTTGAACAAGTACATGACTATATCAGGA 675  
QY 673 GAGGAATTTCTCAAGGTCTTGAAGAGATGATGACAAAGTTCTTGCACTTTTG 732  
DB 676 TTTCAAATTTGAGAAACCATTTGACC--GCGAGACTTCTGGCAATTTAGACAACTACTC 732  
QY 733 AGGTCAACCATTCAGTGTCTGACAAGACCAAGCTTACTTTGTGATGTTCTTCGTCA 792  
DB 733 CTGCTGTGTGTAATCTATTTGAGATATCTGCTTACTTGCAGAGACCTCTATATAT 792  
QY 793 GCAATCAACAAAATGGAATGATGAAGAGCACTCACTAGAAATTTGTACACACAAGAGCT 852  
DB 793 GCTATGAGAGGAGCTGGAGACAGATGATCATACCTCATCAGAGTCAATGTTTCAGAGAT 852  
QY 853 GAGATTGACTTGAAGTCAATTTGAGAGAGTACAGGCGCAGAACAGCAATTCCTTTGAG 912  
DB 853 GAGATTGATCTGTTTAACTCAAGAGAGATTTAGAGAAATTTTGCACCTCTCTTTAT 912  
QY 913 AAGCTATTACAAAGCACTCGTGAATTAAGAGAAATGCTCTGCGACTTCGGGT 972

DB 913 TCCATGATTAAGGAGATATCATCTGAGGACATTAAGAAAGCTCTTCTGCTGCTCCGGA 972  
QY 973 GAAAGTGA 980  
DB 973 GAAAGTGA 980

Search completed: August 22, 2005, 03:09:22  
Job time : 260 secs



OY	121	AACGAGCACTTGATCATATCAATCTTGGCTCACAGAAGTCGTGAACAAGAAACTCATC	180
Dd	121	AACGAGCACTTGATCATATCAATCTTGGCTCACAGAAGTCGTGAACAAGAAACTCATC	180
OY	181	AGCGAAGCATACCAAGAACCTTAGCGCGAAGACTTTCTCAAGACTCTTGAACAAGAGCTC	240
Dd	181	AGCGAAGCATACCAAGAACCTTAGCGCGAAGACTTTCTCAAGACTCTTGAACAAGAGCTC	240
OY	241	TCTAACAATTTCAGAGAGCTATCTTGTGTGGACTCTTGAACCCGGTAGAGCGTAGTCT	300
Dd	241	TCTAACAATTTCAGAGAGCTATCTTGTGTGGACTCTTGAACCCGGTAGAGCGTAGTCT	300
OY	301	TTATTGGCTAATGAAGCTPACAAAAAGATGGAATTGAAGCAACCAAGTTCTTATGGAATT	360
Dd	301	TTATTGGCTAATGAAGCTPACAAAAAGATGGAATTGAAGCAACCAAGTTCTTATGGAATT	360
OY	361	GCTTSCACAAGSACATCAACCGAGCTGTACCGTAGGCAAGCTTACACATCGTCGCATC	420
Dd	361	GCTTSCACAAGSACATCAACCGAGCTGTACCGTAGGCAAGCTTACACATCGTCGCATC	420
OY	421	AAGAAGCTCTTTGAAGAGAGCGTTGCTCACACACTACCGGTGACTTCAGAAAAGCTTTTG	480
Dd	421	AAGAAGCTCTTTGAAGAGAGCGTTGCTCACACACTACCGGTGACTTCAGAAAAGCTTTTG	480
OY	481	GTTTTCTTTGTTAACCCTCATACAGATGCAAGAGATGAAATGAAATGACATGACATTTGCTTAAG	540
Dd	481	GTTTTCTTTGTTAACCCTCATACAGATGCAAGAGATGAAATGAAATGACATGACATTTGCTTAAG	540
OY	541	CAAGAAAGCTAAGCTGTGTCATGAGAAAAATCAAAGSACAGCACTACAAATGATGAGATGTT	600
Dd	541	CAAGAAAGCTAAGCTGTGTCATGAGAAAAATCAAAGSACAGCACTACAAATGATGAGATGTT	600
OY	601	ATTAGAATCTTGTCCACAAAGAGCAAGAGCTCAGATCAATGCTACTTTTAAACGTTAACCA	660
Dd	601	ATTAGAATCTTGTCCACAAAGAGCAAGAGCTCAGATCAATGCTACTTTTAAACGTTAACCA	660
OY	661	CATGATCATATGGCAGAGAAATTTCAAAGACTCTTGAAGAAAGAGATGATGACAAATTC	720
Dd	661	CATGATCATATGGCAGAGAAATTTCAAAGACTCTTGAAGAAAGAGATGATGACAAATTC	720
OY	721	CTTGACACTTTGAAGGTCAACCTTAGAGCTTGACAACAACGACCTTACCTTGTGCAT	780
Dd	721	CTTGACACTTTGAAGGTCAACCTTAGAGCTTGACAACAACGACCTTACCTTGTGCAT	780
OY	781	GTTCTTCGTTGAGCAATCAACAAAAGCTGGAAGCTGATGAGAGAGCACTCACTAGATTTGTG	840
Dd	781	GTTCTTCGTTGAGCAATCAACAAAAGCTGGAAGCTGATGAGAGAGCACTCACTAGATTTGTG	840
OY	841	ACCACAAGAGCTGAGATGACTTGAAGGTCATTTGAGAGAGAGTACACAGCGCAGAAACAGC	900
Dd	841	ACCACAAGAGCTGAGATGACTTGAAGGTCATTTGAGAGAGAGTACACAGCGCAGAAACAGC	900
OY	901	ATTCTTTGGAGAAGCTATTAACCAABAAGCTCGTGAAGATTAAGAAAGATGCTCGTC	960
Dd	901	ATTCTTTGGAGAAGCTATTAACCAABAAGCTCGTGAAGATTAAGAAAGATGCTCGTC	960
OY	961	GCACTTCTCGGTAAGATGATGCTTAAATCAATCAATCTCTCCAAGAGAAACATAAGCTGC	1020
Dd	961	GCACTTCTCGGTAAGATGATGCTTAAATCAATCAATCTCTCCAAGAGAAACATAAGCTGC	1020
OY	1021	TCTACAGCTTCTGTTATCTTTATCTCTATCTCTCTCTCTCTTGTGATGAGTTCAAATCGTT	1080
Dd	1021	TCTACAGCTTCTGTTATCTTTATCTCTATCTCTCTCTCTCTCTTGTGATGAGTTCAAATCGTT	1080
OY	1081	TGATTTTGTGTTCTACAAAAAAGCTTGTGTTTCTGTGTGTGTTTGAAGTTCTTAAATPA	1140
Dd	1081	TGATTTTGTGTTCTACAAAAAAGCTTGTGTTTCTGTGTGTGTTTGAAGTTCTTAAATPA	1140
OY	1141	TGCAAAAAGAGAGACAGAGAGAACCAAGTGTGCTCTTAAGTTATATATATATGAAAG	1200
Dd	1141	TGCAAAAAGAGAGACAGAGAGAACCAAGTGTGCTCTTAAGTTATATATATATGAAAG	1200
OY	1201	CATTGGCTTAAAAAAAAAAAAAAAAAAAAA 1230	

[illegible]

Db	361	TCACGAGAGCTGTTCAACGTCAGGCAAGCTTACATGCTCCGTACAGAAATCTCTTGAA	420
Qy	436	GAGAGCGTTGCTCAACAACCTACCGGTGCTTCAGAAAGCTTTGGTTTCTCTGTACC	495
Db	421	GAGAGCGTTGCTCAACAACCTACCGGTGCTTCAGAAAGCTTTGGTTTCTCTGTACC	480
Qy	496	TCATTACAGGTACGAGAGAGATGATGAAATGATGGCTTAAGCAAGAGCTAAAGCTG	555
Db	481	TCATTACAGGTACGAGAGAGATGATGAAATGATGGCTTAAGCAAGAGCTAAAGCTG	540
Qy	556	GTCCTATGAGAAAATCAAGGACAAGCATGATGATGAGATGTTATTAGATCTTGTCC	615
Db	541	GTCCTATGAGAAAATCAAGGACAAGCATGATGATGAGATGTTATTAGATCTTGTCC	600
Qy	616	ACAAGAGCAAAAGCTAGATCAATGCTACTTTTAA	650
Db	601	ACAAGAGCAAAAGCTAGATCAATGCTACTTTTAA	635

```

RESULT 3
US-10-767-795-4242
; Sequence 4242, Application US/10767795
; Publication No. US2004018130A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Cao, Yongwei
; APPLICANT: Zhou, Yihua
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53534)B
; CURRENT FILING DATE: 2004-01-30
; NUMBER OF SEQ ID NOS: 117596
; SEQ ID NO 4242
; LENGTH: 1184
; TYPE: DNA
; ORGANISM: Gossypium hirsutum
; FEATURE:
; OTHER INFORMATION: Clone ID: GOSHI-09MAY01-C101_1
; US-10-767-795-4242

```

Query Match	40.5%	Score 497.8	DB 19	Length 1184
Best Local Similarity	70.7%	Pred. No. 2.3e-130		
Matches 691	Conservative 0	Mismatches 282	Indels 4	Gaps 2
QY	12	AAACCTAAAAGTAGAGAA-AAATGGCGACTTTAAGTTCCTGATTCCTGCTC	70	
Db	47	AAAGAAATAAAGGAAGAAAGCAATGGCACTCTTACAGTCCCAAGCAATTCCTCG	106	
QY	71	CTTCTGATGATGCTGAGCAATTGAGAACCGCTTTGAAGATGGGGTACGAACGAGCT	130	
Db	107	TGCTGAAGATTGGAAACAGCTAAGAAAGCCTTTTCAGATGGGGAACTAATAGGGCT	166	
QY	131	TGATCATATCAATCTTGGCTCAGAGAGTGTGTAACAGAGAAAGTACATGAGCAAGAT	190	
Db	167	TAACTAATGATATATTGGGTACAGAAATGCCAGACAGAACTTGATTGAAAACCT	226	
QY	191	ACCACGAACCTTACGGCGAAGACCTTCTCAAGACTTTGACAGAGCTCTTAACGATT	250	
Db	227	ACCGTGAACCTTAGAGAGAGATCTCTCAAGGACACTAGACAGAGCTTCGATGACT	286	
QY	251	TTCGAGAGAGCATTTGTGAGACTTTGAACCCGGAGACGTGATCTTATTGGCTA	310	
Db	287	TTGAGAGGCTGGTTTGTCTTGGGCTCTGATCTGCTGAACGATGATGCCCTTTTGGCTA	346	
QY	311	ATGAGCTACAAAAGATGACATTCGAAGCAACAAAGTTCATTATGAAGTTCTTGCAAA	370	
Db	347	ATGAAGCACCAAAAGGTGACCTTCAGCAATCAAGTCTTATGAAATACTTGACAA	406	
QY	371	GGACATCAACGCACTGCTTACGGTAGGCAAGCTTACATGCTCGCTACAGAAAGTCTC	430	
Db	407	GGTCTGCAACCAACTCTTACGCAAGGACGGCTTATCAAGTCTCGTTATTAAGAAAGTCG	466	

Oy	431	TTGAAAGAGAGCTTGGCTCTCACCACTACCGGTGACTTTCAGAAAGCTTTGGTTTCTCTTG	490
Db	467	TTGAAAGGAGATGGTTGCTTCATCCACGACTGGCCGCTTCGGTAAAGCTCTCTTCACTTCTTAG	526
Oy	491	TTACCTCATACAGGTACGGAAGGAGATGAAAGTGAACATGACATTTGGCTTAAGCAAGAGCTA	550
Db	527	TGAAGTCAATACAGATATGAGGAGAGAGAGGTGAACATGAATCTGGCCGAAAAACAGGGCGA	586
Oy	551	AGCTGGTCCATGAGAAAAATCAGGACACAGCACTACAAATGATGAGATGTTATTAAGAACT	610
Db	587	AGTTGCTTCATGAGAAAAATTTACAGCAAGCTTACAGTATGACGATGTCAATAGGGGTTT	646
Oy	611	TGTCCACAAGAGGAAAGCTCAGATCAATGCTACTTTTAACCGTTACCAAGATGATCATG	670
Db	647	TGGCTAACAGAGGACAGGACAGATCAATGAACTTCGAAATCACTAACAAAAATGAATATG	706
Oy	671	GCGAGAAATTTCTCAAGATCTTGAGAGAGAGATGATGATGACAGTTCCTTGACATTT	730
Db	707	GAAATGACATTAACAGAGACTT---GAAAGCTGATCCTTAAGATGAAGTTCTTGACACTAC	763
Oy	731	TGAGGTCAACCATTTCAAGTCTTGAACAAGCCAGAGCTTTATCTTTGCGATGTTCTTGCTT	790
Db	764	TAAAGTCCACAGTGAAGTGGCTGGTTCATCCGAAAAAGTATTTTGAAGAAGTTCTTGCC	823
Oy	791	CAGCAATCAACAAACCTGGAACATGATGAAGAGACCTCACTAGCAATTTGTGACCAAGAG	850
Db	824	TAGCAATCAATATAGCGAGAACCGATTAAGAGAGCTCTTACTAGAGTTGTTGCACTAGGG	883
Oy	851	CTGAGATTGACTTGGAAGGTCAATGGAGAGATACAGCCAGAGAACAGCAATTCCTTAGG	910
Db	884	CTGAGGTTGATCTTAAGATCATATGACAGATGATGATACAGCCAGAGAACAGTGTCCCACTGA	943
Oy	911	AGAAAGCTATTTACAAAGACACTTGTTGAGATTAACGAAAGATCTGTCGCACTTCTCG	970
Db	944	CTCGTGCCATTTGTCAAAGACACTTCATGGAAGCTATGAATAATTCCTGCTGTACTTGCAAG	1003
Oy	971	GTGAAGATGATGCTTAA 987	
Db	1004	GACATGTGAGAAATGA 1020	

```

RESULT 4
US-10-425-114-7439
/ Sequence 7439, Application US/10425114
/ Publication No. US2004003488A1
/ GENERAL INFORMATION:
/ APPLICANT: Liu, Jindong
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Screen, Steven E
/ APPLICANT: Tabaska, Jack E
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
/ TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
/ FILE REFERENCE: 38-21(53313)B
/ CURRENT APPLICATION NUMBER: US/10/425,114
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 73128
/ SEQ ID NO 7439
/ LENGTH: 1145
/ TYPE: DNA
/ ORGANISM: Glycine max
/ FEATURE:
/ OTHER INFORMATION: Clone ID: 700661911_FTI
US-10-425-114-7439

```

Query Match	39.0%	Score 479.6	DB 18	Length 1145
Best Local Similarity	69.0%	Pred. No. 3.3e-125		
Matches 672	Conservative	0	Mismatches 299	Indels 3
			Gaps	1

  

14	ACACTAAAGTGAAGAAAAATGCGCACTCTTAAGCTTTCTGATCTTGTTCTCTCTT	73
34	AAACAAAAACGAAACACATGCGCACACTTAAGTCTCTGAGCTCTTCTCCCGCGTTG	93

Oy	7	CTGATGATGCTGAGCAATTGAGACCGCTTTTGAAGATGGGGTATCGAAGAGGACTTGA	133
Db	94	CAGATGATGTGTGAGAGCTCCAGAAAAGCTTCTCAGTTTGGGAACTTAAACGAGAGCTGA	153
Oy	134	TCATATCAATCTTGGGCTCAACAGAGTGTGAAACAGAGGAAAGTCATCAGCCAGACATACC	193
Db	154	TTGTATGATCTTGGCTGTCAATAGGAATGCTGCTCAGAGGAGCTTATACGAAACTTATG	213
Oy	194	ACGAAACTTACGGCGAAAGACTTCTCAGACTCTTGAAGAGAGCTCTAACATTTTCG	253
Db	214	CCAGAGACTTATGGAAGAAAGATCTCTCAAGCCCTTGAGACAAAGAACTCAGAGTATTTTG	273
Oy	254	AGAAGCTATCTTGTGTGTGGACCTTGAACCGGGTGAAGGTGATGCTTATTTGGCTATG	313
Db	274	AGAGGCTGTTCATCTTTGGACCTTGATTTCTGTGAACGTATGCAATTTTGGCAACG	333
Oy	314	AAGCTACAAAAGATGAGACTTCAAGCAACCAAGTCTTATGGAAGTTGCTTGACAAAGA	373
Db	334	AGGGACTTAAATAATGACTTCAAGCAATCAGGTTCTGTGTGAAATAGCTCGCACTAGGT	393
Oy	374	CATCAACGACGCTGCTTCAAGCTTAGGCAAGCTTACCATGCTTCGCTACAAAGATCTCTTG	433
Db	394	CCCTGGAACAATTTGTTGCTGCGAGGAAAGGCTTACCAATGTTCTTATTAAGAAAGTCTGG	453
Oy	434	AAGAGACTTGTGCTCAACACACTACCGGTGACTTTGAGAAAGCTTTTGGTTCTCTTTTA	493
Db	454	AGGAGATGTTGTCTATCAACACAAAGAGACTTTCGTAAGTCACTACTACTCTTGATTA	513
Oy	494	CCTCATACAGGTACGAAGAGATGATGAATGATGACATTTGGCTTACGAAAGACTTAAGC	553
Db	514	GTTCTTTATGATATGAAGAGATGAGGTCAATTGACCTGGCAAAAATGAGGCAAAAT	573
Oy	554	TGCTCCATGAAAAATCAAGACACAGCACTACATGATGAGAGTATTGAATCTTGT	613
Db	574	TGCTGCATGAGAAAGATTTCAAACAAAGGCTTATATGATGACGACTTCATCAGGATTTTGG	633
Oy	614	CCAAACAACCAAGCTCGAGATCAATGCTACTTTTAAACGTTACCAAGATGATATGAGCG	673
Db	634	CCAAAGAGACCAAGGACAGATTAATGCTACTTTGAATCACTCAAAATATCATTGTGAC	693
Oy	674	AGGAAATCTCAAGAGCTTGTAGGAAGAGATGATGATGACAAAGTCTCTTGCACTTTTGA	733
Db	694	AAGATATCAACAAAGAGACT---GAGAGCTGATCCAAAGAGACAGTTCCTGCTTTACTTA	750
Oy	734	GGTCAACCAATTCAGTGTCTTGAACAAGCCAGACTTTACTTTGTGAGTTCTTGTCAAG	793
Db	751	GAGCAACTGTGAGTGTGTTGATCCGTCCGAGAAAGTACTTGAGAAAGGTTGTTGCTTGG	810
Oy	794	CAATCAACAAAACCTGGAATGATGAAAGAGCACTCACTATGAATTTGTACCCAAAGACTG	853
Db	811	CGATCAACCAAGCGAGGAAACGATGAAGAGCCCTTTTAAAGAGTGTGTGCCCAAGGGCTG	870
Oy	854	AGATTGACTTGAAGGTCAATGTGAGAGAGATACAGAGCGAGAAACGACTTCTTTGGAGA	913
Db	871	AGGTGATTTGGAAAAACATAGCAGATGATATCAAAAGAGAGCAGTCCCTCTTTGAC	930
Oy	914	AAGCTATTTACAAAGACACTCGTGAAGATTACGAGAAAGTCTGCTGCACTTCTCGGTG	973
Db	931	GTCCTATTGTTTAAGACACACACCGGTGACTATGAGAAATGCTTGTGCGCTTTTAAAGAC	990
Oy	974	AAGATGATGCTTAA 987	
Db	991	ATGATGATGCTTGA 1004	

```

RESULT 5
US-10-425-114-29928
; Sequence 29928, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua

```

```

?  APPLICANT: Kovalic, David K.
?  APPLICANT: Screen, Steven E
?  APPLICANT: Tabaska, Jack E
?  APPLICANT: Cao, Yongwei
?  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
?  TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
?  FILE REFERENCE: 38-21(5313)B
?  CURRENT APPLICATION NUMBER: US/10/425,114
?  CURRENT FILING DATE: 2003-04-28
?  NUMBER OF SEQ ID NOS: 73128
?  SEQ ID NO 29928
?  LENGTH: 1145
?  TYPE: DNA
?  ORGANISM: Glycine max
?  FEATURE:
?  OTHER INFORMATION: Clone ID: UC-GMEOPIC033E01_FLI
US-10-425-114-29928

```

Query Match	39.0%	Score 479.6;	DB 18;	Length 1145;
Best Local Similarity	69.0%;	Pred. No. 3.3e-125;		
Matches 672; Conservative	0;	Mismatches 299;	Indels 3;	Gaps 1,

OY	1	ACACTAAAGTAGAAGAAATATGGGACTCTTAAGCTTTGATTCGTGTCCTGCT	73
Db	34	AAACAAAAACAGAACAAACAATGGGACCTTAAGGTTCTCAGCCTCTTCCCCCGTTG	93
OY	74	CTGATGATGCTGAGCAATTTGAGAACCGCTTTTGAAGATGAGGGTACGAACGAGACTTGA	133
Db	94	CAGATGATTTGGAGGAGCTCCGAAAGGCTTCTCAGGTTTGGGGAATCAACGAGAGCTGA	153
OY	134	TCATATCATCTTGGCTCAACAGAGTCTGAACAGAGAAAGTCAATCAGGCAAGCATACC	193
Db	154	TTGTATCATCTTGGCTCATGAGATGCTGCTCAAGGAGCTTATCAGAGAACTTAAAG	213
OY	194	AOGAAACCTAAGGGGGAAGACCTTCCTCAAGACTCTTGAACAAGGAGCTCTAAGATTGCG	253
Db	214	CCGAGACTTAAGAGAAAGATCTCTCAAGGCTTTGACAAAGAACTCAGAGATGATTTGG	273
OY	254	AGAGAGCTATCTTGTGTGGACTCTTGAACCGGTAAGGCTGATCTTATTTGGCTAATG	313
Db	274	AGAGGCTGTTCATCTTTGGACACTTGAATTCGTGGAAGTATGATTTGGGCAAGCG	333
OY	314	AAGCTACAAAAGATGAGCTTCAAGCAACCAAGTTCTTATGGAAGTTGCTTGCAAGGA	373
Db	334	AGGCACTTAAAAAAATGAGACTTCAAGCAATCAGGTTCTGCTGGAATAAGCTCAGCATGCT	393
OY	374	CATCAAGCAGCTGTTCAAGCTAGGGAAGCTTACATGCTGCTACAAAGAGTCTCTTG	433
Db	394	CCTCTGAACAAATGTTGCTCTGGAGGAAGGCTTACCAATGTTTATATAGAAGATCTCTGG	453
OY	434	AAGAGGAGGTTGCTACCACTACCGGTGACTTCAGAAAGCTTTGTTTCTCTGTGTA	493
Db	454	AGGAGGATGTTGCTCATACACAAACAGAGACTTCCGTAAAGCTCATCTACTCTGGTTA	513
OY	494	CCTCATACAGGTACGAGAGAGATGAATGAACATGACATTTGGCTTAAACAAGAGCTTAAC	553
Db	514	GTTCTTATCATATATGAAGAGATGAGTCAACTTGACCCCTGGCAAAAATTGAGGCAAAAT	573
OY	554	TGCTCCATGAGAAAATCAAGACAAGCACTCAATGAAGAGATGTTATTAATCTGT	613
Db	574	TGCTGCATGAGAAAGTTTCAACCAAGGCTTATTAATGAATGACACTTCACAGATTTTGG	633
OY	614	CCACAAGAGCAAAAGCTCAGATCAATGCTACTTTTAAACGTTTACCAGATGATCATGGCG	673
Db	634	CCACAAGAGCAGAGACACAGATTAATGCTACTTTGAATCATTCAAAAGATGCAITTTGGAC	693
OY	674	AGGAATTTCTCAAGAGCTTTGAGAGAGAGATGATGATGACAAGTTCTTGAACATTTTGA	733
Db	694	AAGATATCAACAAGGACCT---GAGGCTGATCAAAAGACGAGTTCTGTCTTACTAA	750
OY	734	GGTCAACATTCAGTGTCTTGAACAGACCAAGAGCTTACTTTGTGCATGTTCTTCTGTCAG	793
Db	751	GAGCAACTGTGAAGGCTTGAATCCGCTCCGGAAGAGTACTTCAGAAAGGTTGTTCGCTTGG	810

QY 794 CAATCAACAAACCTGSACTGATGAGAGCACTCACTAAGATTGTGACCAAGACTG 853  
DB 811 CGATCAACAGCGAGAAACCGATAGAGAGCCCTTACAAAGAGTGTCACCAAGGCGTG 870  
QY 854 AGATTGACTTGAAGTCACTGAGAGAGTACCAAGCGAAGAAAGCACTTCTTGGAGA 913  
DB 871 AGTTGATTTGAAAACATAGACATAGATTCAAAGAGAGAGAGATGTCTCTTGAAC 930  
QY 914 AAGCTATTACCAAGACACTGCTGAGATTACGAAAGATCTGTCGCACTTCTCGTG 973  
DB 931 GTGCTATTGTTAAGACACCAACCGGTGACTATGAGAAAATGCTTGTGGCGCTTTAGAAC 990  
QY 974 AAGTGAATGCTTAA 987  
DB 991 ATGATGATGCTTGA 1004

## RESULT 6

US-10-424-599-1190  
; Sequence 1190, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 1190  
; LENGTH: 1393  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_101073C.1  
US-10-424-599-1190

Query Match 39.0%; Score 479.6; DB 18; Length 1393;  
Best Local Similarity 69.0%; Pred. No. 3.7e-125;  
Matches 672; Conservative 0; Mismatches 299; Indels 3; Gaps 1;

QY 14 ACACTAAAGTAGAAGAAAATGCGACTTTAAGTTTGTGATTCTGTCTGCTCCT 73  
DB 113 AAACAAAACAGAAACAAATGCGACACTTAAGTTTCTCAGCCTTCCCGCGTTG 172  
QY 74 CTGATGATGCTGAGCAATTGAAACCGCTTTTGAAGATGGGTAAGCAAGAGACTTGA 133  
DB 173 CAGATGATGCTGAGCACTTCCGAAAACCTTCTCAGGTGGGAACTTAAGAGAGCTGA 232  
QY 134 TCATATCAATCTTGGCTCAAGAGATGCTGAACGAGAAAGATCATGAGCAAGCATTAC 193  
DB 233 TTGATGATCTTGGCTCAAGAGATGCTGCTCAGAGAAAGCTTATCAGAAAATTATG 292  
QY 194 ACGAAACCTTACGCGGAAAGACCTTCTCAAGACTTTCAGCAAGAGACTTCTTAACGATTTCG 253  
DB 293 CCCAGACTTATGAGAGATCTCTCAAGGCTTTCAGCAAGAACTCAGCGATGATTTTG 352  
QY 254 AGAGAGCTATCTTGTGTGACTTTGAACCCGATGAGCTGATGCTTATTTGGCTAATG 313  
DB 353 AGAGGCTGTTCATCTTTGACACTTATCTTGTCTGAACGATGCAATTTTGGCAACG 412  
QY 314 AAGCTACAAAAGATGAGACTTCAAGCAACCAAGTTCTTATGAGAGTTGCTTGCACAGGA 373  
DB 413 AGGGAATTAAGAAATGAGCTTCAAGCAATCAGGTTTCTGAGAAATGCTTGCACAGGT 472  
QY 374 CATCAAGCAGCTCTTCAAGCTTGAAGCAAGCTTACATGCTCGCTACAGAGTCTTGTG 433  
DB 473 CCTCTGAACAATGTTTGTGCTGCGAGGAAGGCTTACCAATGTTCTTTATAGAAGTCTCTGG 532

QY 434 AAGAGAGCTTGTCTCAACCACTACCGGTGACTTCAGAAAGCTTTTGTCTCTGTTA 493  
DB 533 AAGAGAGTGTGTCTCATCAACAACAGAGACTTCGTAAGCTTCATCTACTCTTGTGA 592  
QY 494 CCTCATACAGGTACGAAAGAGATGAAAGTGAACATGACATTTGGTTAAGCAAGACTTACG 553  
DB 593 GTTCTTATGATATGAAAGAGATGAGGTCACTTGAACCTCGGAAAACATGAGCAAAAT 652  
QY 554 TGGTCATGAGAAATATCAAGCAAGCACTACATATGATGAGATGTTATTAAGATCTTGT 613  
DB 653 TGCTGATGAGAGATTTCAACAGGCTTATATATGATGACACTTCATCAGATTTTGG 712  
QY 614 CCAAGAAAGCAAGCTCAGATCAATGCTACTTTTAACGTTTCAAGATGATCATGAGCG 673  
DB 713 CCAAGAGAGCAAGCAAGATTAATGCTACTTTTAATCACTCAAGATGCAATTTGGAC 772  
QY 674 AAGAAATTCACAGAGCTTGAAGAGAGATGATGACAAAGTTCCTTGCATTTTGA 733  
DB 773 AAGATATCAACAAAGAGACT---GAGGCTGATCCAAAGAGAGAGTTCCTTACTTA 829  
QY 734 GTTCAACATTCAGTCTTGAACAAGCAAGACTTATCTTGTGATGTTCTTGTGAG 793  
DB 830 GAGCAACTGTGAGTCTTGTATCCGTCGAGAGAGTACTTCGAAAGTGTTCGCTTGG 889  
QY 794 CAATCAACAAACCTGGAATGATGAGAGCACTCACTGAATTGTGACCAAGAGCTG 853  
DB 890 CGATCAACAAAGCAAGAGACGATGAAAGAGCCCTTCAAGAGGTGTGCCCAAGAGCTG 949  
QY 854 AGATTGACTTGAAGTCACTTGAAGAGAGTACCAAGCGCAAGCAACATCTTGTGAGA 913  
DB 950 AGTTGATTTGAAAACATAGCAAGATGATGATTAAGAGAGAGACATGTCTCTTGAAC 1009  
QY 914 AAGCTATTACCAAGCACTCTGTGAGATTACGAAAGATGCTGTGCACTTCTCGTG 973  
DB 1010 GTGCTATTGTTAAGACACCAACCGGTGACTATGAGAAAATGCTTGTGGCGCTTTAGAAC 1069  
QY 974 AAGTGAATGCTTAA 987  
DB 1070 ATGATGATGCTTGA 1083

## RESULT 7

US-10-219-220-218  
; Sequence 218, Application US/10219220  
; Publication No. US20030082724A1  
; GENERAL INFORMATION:  
; APPLICANT: Film, Barry  
; APPLICANT: Lasham, Annette  
; TITLE OF INVENTION: Compositions affecting programmed cell  
; FILE REFERENCE: 11000.1022c1  
; CURRENT APPLICATION NUMBER: US/10/219,220  
; PRIOR FILING DATE: 2002-08-14  
; PRIOR APPLICATION NUMBER: U.S. No. US20030082724A1 09/325,932  
; PRIOR FILING DATE: 1999-06-04  
; NUMBER OF SEQ ID NOS: 290  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 218  
; LENGTH: 1058  
; TYPE: DNA  
; ORGANISM: Eucalyptus grandis  
US-10-219-220-218

Query Match 38.5%; Score 473.4; DB 14; Length 1058;  
Best Local Similarity 69.0%; Pred. No. 1.8e-123;  
Matches 664; Conservative 0; Mismatches 296; Indels 3; Gaps 1;

QY 25 AGAAGAAAATGCGACTCTTAAAGTTTGTGATTTCTGCTGCTCTTCTGATGATGCT 84  
DB 63 AGAGAGAGATGCGACTATCGCGGTCCACCGTCTGCTTCCGCTCGGCTGAGAGTGC 122  
QY 85 GAGCAATTTGAAACCGCTTTTGAAGATGGGGTACGAGAGACTTGAATCATCATTC 144



Db 123 GAGCAGCTCCAAAAAGCTTTGCGAGAGTGGGAAACGAATGAATCATATCATATCATA 182  
Qy 145 TTGGCTCACAAGAGTGTGAACAGAGAAAGTATCAGGCAACCATATCCAGAACTTAC 204  
Db 183 CTGGCTCACAAGAAACGAGCGCAGCGAAAGTATCCGAAACATATGCGGAGACATAT 242  
Qy 205 GCGGAAGACCTTCTCAAGACTTCTGACAAAGAGCTCTTAACGATTTCCAGAGAGCTATC 264  
Db 243 GGGGAAGATCTTCTCAAGAGCTTGAACAGAACTCTTAAGTACTTTGAGAGATCTGTC 302  
Qy 265 TTGTTGTGACTTTGAACCCGGTGAAGCTGATGCTTTATTTGGCTAATGAAGCTTACAAA 324  
Db 303 CTCTGTGGACCTCTGGATCTCGGAGCGGTGATGCAATTTCTTGCATGAAGCTTACCAAG 362  
Qy 325 AGATGACCTTCAAGCAACGAATTTCTTAATGGAAGTTGCTTGCACAGGACATCAAGCGAC 384  
Db 363 AGATGACTTTCAGCAACTGGGTTCTCATGAAATGCTTGCACAGAGCTTTCATATGAGAG 422  
Qy 385 CTGCTTCACGCTAGGCAAGCTTACCAATGCTGCTACAGAAAGTCTCTTGAAGAGAGCTT 444  
Db 423 TTATTCATGCTGAGGCGAGGCTTATCATGCTCGTTATAGAAATCTCTTGAAGAGAGATC 482  
Qy 445 GCTTCACCACTACCGGTGACTTTCAGAAAGCTTTGGTTTCTTTGTATCTTATACATACAG 504  
Db 483 GCATATCACACTACCTGGGATTTCCGCAAGCTCTTGTCTCTGGCAAGTACCTTTCGG 542  
Qy 505 TAGAAGAGGATGAAGGACATGACATTTGGCTAAGCAAGAACTAAGCTGCTGCTAGAG 564  
Db 543 TATGAGGGGCTTGAAGGACATGACATTTGGCGAGATCAGAGGCTTAAGATCTTCAATGAG 602  
Qy 565 AAAATCAAGGACAGCACTCAATGATGAGATGTTATTTAGATCTTTCACAGAAAGC 624  
Db 603 AAGATTCACAGAAAGGCTTACATCATGATGAGCTCATCGAATTTGTACTACAGAAAGT 662  
Qy 625 AAAAGCTCAGTCAATGCTTCTTAAACCTTACCAAGATGATCATGGCGAGAAATCTTC 684  
Db 663 AAAAGCTCAGTCAATGCTTCTTAAACCTTACCAAGATGATTTGGGAATGCCATCAAC 722  
Qy 685 AAAAGCTTGAAGAGGAGATGATGATGACATTTCTTGCATCTTCTTGAAGTCAACAT 744  
Db 723 AAGGATCT--GAAAGCTGATCCAAATGATGATTTCTGAATGCTGAGATCAGCAAT 779  
Qy 745 CAGTGTCTGACAAAGCAGAGCTTCTTCTTGTGATGCTTCTGCTTCAAGCAATCAACAA 804  
Db 780 AAGGCTTGACTTATCCAGAGAAAGTACTTTGAAGAGTACTGCGTCCGCAATCAACAA 839  
Qy 805 ACTGGAATCATGATGAGAGCACTCATGATTTGTGACCAAGAGCTGATGATTCCTG 864  
Db 840 CTGGGAACAGATGAATGGGCTCTTTCACAGATGATCCACTGCTGGGGAAGTTGACATG 899  
Qy 865 AAGGTCAATGAGAGGAGTACACGCGCAGGAACAGCAATTCCTTGGAGAAAGCTATTAAC 924  
Db 900 CACCGGATCAAAAGACATGACACAAAGAGAAACAGTGTCTCTAGATCGGGCTATTTGCA 959  
Qy 925 AAAAGCACTGCTGAGATTTACGAGAAAGATCTGCTGCACTTCTCGGTGAAGATGATCT 984  
Db 960 AATGACACATCTGGGAGCTTACAGAAAGATCTTCTGCTTTGGTTGGAACATGAGATGCT 1019  
Qy 985 TAA 987  
Db 1020 TGA 1022

RESULT 8  
US-10-393-840-28  
; Sequence 28, Application US/10393840  
; Publication No. US2003022922A1  
; GENERAL INFORMATION:  
; APPLICANT: Bloksberg, Leonard N.  
; TITLE OF INVENTION: Materials and Methods for the  
; FILE REFERENCE: 11000.1012c3  
; CURRENT APPLICATION NUMBER: US/10/393,840

; CURRENT FILING DATE: 2003-03-20  
; PRIOR APPLICATION NUMBER: US 09/636,800  
; PRIOR FILING DATE: 2000-08-10  
; PRIOR APPLICATION NUMBER: US 09/170,862  
; PRIOR FILING DATE: 1998-10-13  
; PRIOR APPLICATION NUMBER: US 60/148,426  
; PRIOR FILING DATE: 1999-08-11  
; PRIOR APPLICATION NUMBER: PCT NZ/99/00169  
; PRIOR FILING DATE: 1998-10-08  
; NUMBER OF SEQ ID NOS: 956  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 28  
; LENGTH: 2588  
; TYPE: DNA  
; ORGANISM: *Eucalyptus grandis*  
US-10-393-840-28

Query Match 38.4%; Score 471.8; DB 17; Length 2588;  
Best Local Similarity 68.6%; Pred. No. 8.7e-123;  
Matches 666; Conservative 0; Mismatches 302; Indels 3; Gaps 1;

Qy 30 AAAATGGGACTCTTAAGTTTCTGATTTCTGCTCTCTTGTGATGATGCTGAGCA 89  
Db 1386 AACAAATGCTGACTTCAACCGTCCGACGACTGCCCCCTGAGCGGATGCTGGAGCA 1445  
Qy 90 ATTGGAACCGCTTTTGAAGATGGGGTACGAACGAGACTTATCATATCTTGGC 149  
Db 1446 GCTCCGAGACGCTTTCGAGAGATGGGGAACAAATGAAAGCTATCTATTCATTTGGG 1505  
Qy 150 TCACAGAAATGCTGAACAGAGAAATCATCAGGCAAGATCAACGAATCTTACGGCGA 209  
Db 1506 TATAGGAATGCGGGCGACAGAGAGCTGATTCGGCAACCTATGCGAGACTTACGGCGA 1565  
Qy 210 AGACCTTCTCAAGACTCTTGAACAGAGCTCTTAAACATTTTCGAGAGATCTCTGT 269  
Db 1566 GGAACCTCTCAAGAGATTTGACAGAGAACTTACCAATGATTTTCGAGAGCTGTGCTCT 1625  
Qy 270 GTGGACTCTTGAACCGGAGCTGATGCTTATTTGGCTAATGAAGTACAAAGATG 329  
Db 1626 TTGATCTCTGATTCGGCTGAGAGCTGATGCTGATCTTGGGATGAGAGCAAGAAATG 1685  
Qy 330 GACTTCAGCAACCAAGTTCTTATGAAATGCTTTCACAGAGATCAACGCACTGCT 389  
Db 1686 GACTTCAGCAACCAAGTTCTTATGAAATGCTTTCACAGAGATCAACGCACTGCT 1745  
Qy 390 TCAGCTAGGCAAGCTTACATGCTCGTACAAAGATCTCTTGAAGAGAGAGTGTCTCA 449  
Db 1746 TATGCAAGACAGCAATATCATGCTCCGATCAAGAAATGCTGGAAGAGAGCTGCTCA 1805  
Qy 450 CCACACTACCGGTGACTTCAAGAAAGCTTTGTTTCTCTTGTATCTCATACAGTACGA 509  
Db 1806 CCACCAATGAGATTTTCGTAAGTGTGCTGATCTCTTGTGAGCTCTCATACGTTATGA 1865  
Qy 510 AGGAGATGAATGAACATGACATTTGGCTAAGCAAGAAAGCTAAGCTGCTCATGAGAAAT 569  
Db 1866 TGAAGATGAGTGAATATGACTTTGGCAAAACAGAGGCTAAGATATCTCAAGAAAGAT 1925  
Qy 570 CAAGACAGCACTACATGATGAGATGTTATTTGAATCTTGTCCACAAAGAGCAAGC 629  
Db 1926 CTCAAGAGAGGCTTATGCGCATGAGATCTCAAAAGATTTTGGCTACTAGAGCAAGC 1985  
Qy 630 TCAGATCAATGCTACTTTTAAACGTTAACCAAGATGATCATGCGAGAAATCTCAAGAG 689  
Db 1986 ACHGTCATGCTTACGCTGATACATCAACAAATGATTTGAAATGATATCAACAGAGA 2045  
Qy 690 TCTTGAAGAGAGATGATGATGACAGATTTCTTGTGACTTTTGAAGTCAACATTCAGTG 749  
Db 2046 TTT--GAAATCTGATCCAAAGAGCGGTTCTTACTATATACGAGAGCTACAGTAAAGTG 2102  
Qy 750 CTTCAGCAAGACAGAGCTTACTTTGTGATGATTTCTTGTGTAGCAATTAACAAACTGG 809  
Db 2103 CCTGACTCCCTGAGAAAGTATTTGAAAGGTTCTTGTCTTACATCAATTAAGCGAGG 2162



```

Oy 450 CCACACTACCGCGTACTCTACAGAAAAGCTTTGGTTTCTCTGGTACCTCATACAGTACGA 503
Db 511 CCACACCAACTGAGATTTTTCGTAAAGTGTCTGGTACTCTTTGGAGCTCTACCGTAATGA 570
Oy 510 AGAGATGAATGAACATGACATGCGCTACAGCAAGAACTAAAGCTGGCTCATGAGAAAT 569
Db 571 TGGAGATAGGTGAATATGACTTTGGCAAAAAGCAGAGGCTTAAGATCTCCACGAGAAAT 630
Oy 570 CAAGGACAAGCACTACAAATGATGAGAGATGTTATTGAATCTTGTCCACAAGAACAAAGC 629
Db 631 CTCAGAGAAAGGCTTAATGGCCATGAGAGATCTCATTAAGGATTTTGGCTTACGAGCAAGC 690
Oy 630 TCAGATCAATGCTACTTTTAACCGTTACCAAGATGATCATGCGCAGGAAATTTCCAAAG 689
Db 691 AAGGCTCAATGCTACGCTGAATCTACTACAAAATGATGTTGGAAATGATTCACAAAGGA 750
Oy 751 TTT---GAAAATGATCCAAAAGAGCGGTTCTCTTACTATGAGAGCTACAGTAAAGTG 807
Db 750 CTTGACAAAGACCAAGCTTTTACTTTGTGCAATGTTCTTGTTCAGCAATCAACAAACTGG 809
Oy 808 CCGTACTCGCCCTGGAAGTATTTTGAAGAGGTTCTTGTCATGACATCAATTAAGCGAG 867
Oy 810 AACTGATGAAGAGACACTCATGAAATGTGACCAAGAGCTGAGATTTGACTTGAAGGT 869
Db 868 AACAGATAAAGGGCTCTGACCAAGATGTTGCTACCGAGGCGCAGAGTTGACATGAATTT 927
Oy 870 CATTGAGAGAGAGTACCAAGCCAGAGAACAGCATTCCTTTGAGAAAGCTATTACCAAGA 929
Db 928 TATAAGTAGAGAGTACCAAGAGAGGAATGAGATCCTCTCGATCGTGCAATTTGCAAGA 987
Oy 930 CACTCGTGAATTAACGAGAAATGCTGTCGCACTTCTCGGTGAAGATGATGTTATC 989
Db 988 CACTACTGGAAGACTATGAAAAAATGCTTCTGGCAATTGATGCGCACGTCGAGGCTTGATT 1047

Oy 990 AATCAATCCTC 1000
Db 1048 TACAAGTACTC 1058

RESULT 11
US-10-393-840-109
; Sequence 109, Application US/10393840
; Publication No, US20030229922A1
; GENERAL INFORMATION:
; APPLICANT: Blockberg, Leonard N.
; TITLE OF INVENTION: Materials and Methods for the
; FILE REFERENCE: 11000.1012c3
; CURRENT APPLICATION NUMBER: US/10/393,840
; CURRENT FILING DATE: 2003-03-20
; PRIOR APPLICATION NUMBER: US 09/636,800
; PRIOR FILING DATE: 2000-08-10
; PRIOR APPLICATION NUMBER: US 09/170,862
; PRIOR FILING DATE: 1998-10-13
; PRIOR APPLICATION NUMBER: US 60/148,426
; PRIOR FILING DATE: 1999-08-11
; PRIOR APPLICATION NUMBER: PCT NZ/99/00169
; PRIOR FILING DATE: 1999-10-08
; NUMBER OF SEQ ID NOS: 956
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 109
; LENGTH: 1293
; TYPE: DNA
; ORGANISM: Pinus radiata
; US-10-393-840-109

```

Query Match	38.1%;	Score 468.6;	DB 17;	Length 1293;
Best Local Similarity	68.4%;	Pred. No. 4,8e-122;		
Matches 664;	Conservative 0;	Mismatches 304;	Indels 3;	Gaps 1;

QY	30	AAAAATGGCGACCTCTTAAGGTTTCATGTCCTGTTCCGTGCTCTTCTGATGATGCTGAGCA	89
Db	91	AACAAATGTGACCTCCACCGTCCCGCAGCCACTGCCCCCTGTAGCCGATATACGTGCGAGCA	150
QY	90	ATTGAGAACCGCCTTTTAAAGGATGGGGTATCGAAACGAGGACTTGTATCATATCAATCTTGGC	149
Db	151	GCTCCGGAAGACCTTCCAGGATGGGGAAACAATGAGAACCTGATTCATATCCATATTTGGG	210
QY	150	TCACAGAAATGCTGAACAGAGAAAGTCATCAGGCAAGCATACACGAAACCTTACGGCGCA	209
Db	211	TCATAGGAATGCGCGCAGAGGAAAGCTGATTCGGCAAACTTATGCCGAGACTTACCGCGCA	270
QY	210	AGACCTTCTCAAGCTCTTGAACAAGGAGCTCTCAAGATTTTGGAGAGACTATCTTGGT	269
Db	271	GGACCTCTCCAGGCANTTGGACAGAACTTACCAATGATTTTGGAGAGCGCTGGTGCTCT	330
QY	270	GTGACCTCTTGAAACCCGATGAGCGCTGATGCTTTTATTTGGCTTATATGAACTTCAAAAAGATG	329
Db	331	TTTGGTCACTTGAATCCGGCTGAACGTGATTTGGTACTTTGGCGAAATGAAGCGACGAAAAGATG	390
QY	330	GACTTCAAGCACCAAGTTCTTATGAAAGTTGCTTGGACAAGAATCATCAACGCAAGCTGCT	389
Db	391	GACTTCAAGCACCAAGGTTCTCAAGAAATAGCTTGACAGAGTCTCCGAGCAGATTTGCT	450
QY	390	TCACGCTAAGGCATTTACCATGCTGCTGCTCAAGAACTCTCTTGAAGAGACGTTGCTCA	449
Db	451	TATGGCAAGACAGACATATCATGCTCCCATACMAAGTCAATGGAAAGGACGTCCGCTCA	510
QY	450	CCACACTACCGGTGACTTCAGAAAAGCTTTGGTTCTCTTGTATTACTCATACAGTTACGA	509
Db	511	CCACACAACTGGAGATTTTGGTAAAGTTGGTATGCTCTTGGGAGCTCTTACCGTATATGA	570
QY	510	AGGAGTAAAGTGAACATGACATTGGGCTAAGCAAGAAAGCTAAGCTGCATCATAGAGAAAT	569
Db	571	TGGAGATGAGGTGATATGACTTTGGCAAAAGCAGAGGCTTAAGTACTTCCACGAGAAAT	630
QY	570	CAAGGACAGACATACATGATGATGATGTTATTAGATCTTGTCCAAGAAAGCAAAAGC	629
Db	631	CTCAGAGAAAGGCTTATGGCCATGAGATCTCATAGGAATTTTGGCTACTGAGAGCAAAAGC	690
QY	630	TCAGATCATGCTTATTAACCGTTACCAAGATGATCATGGCGAGAAATTTCTCAAGAG	689
Db	691	ACAGGTCAATGCTTACGCTGAATCACTCAAAAATGATTTGGAAATATATATTCACAAGGA	750
QY	690	TCTTGAGGAAGGAAATGATGATGACAGTTCCTTGGCACTTTTGAAGTCAACCATTTGAGT	749
Db	751	TTT--GAAACTATCCAAAAGACGCGTTCCTTATCATACTGAGACTACAGTAAAGTG	807
QY	750	CTTGAACAAGACAGAGCTTTACTTTGTGATGTGTCTTCTGTTACAGCATCAACAAACTGG	809
Db	808	CTGATCTGCCCTTGAAGATTTTGAAAAAGGTTCTTCTGCTAGCATCAATTAAGCAGAG	867
QY	810	AACTGATGAAGAGCACTCACTAGAAATTGTGACCAACAAGCTGAGATTTGACTTTGAAGGT	869
Db	868	AAACGATGAAGGGGCTCTGACCAAGATGATTTGCTACCAAGGGCCGAGGGTTACATGAAGTT	927
QY	870	CATTGGAGGAGATACAGAGCGCAGAAACAGATTTCTTTGAGAAAGCTTATTTACAAAGA	929
Db	928	TATATAGTAGAGATACAGAGAGAGAAATAGCATCTCTCTGATGTGTCCATTTGTCAAGGA	987
QY	930	CACCTGTGAGATTACGAAAGATGCTGTGTGCACTTCTCGGTGAAGATGATGCTTTATC	989
Db	988	CACACTGTGAGACTATGAAAAATGCTTGTGGCATGATGATTTGGCCACGTGAGGCTTGAT	1047
QY	990	AATCAATCTTC 1000	
Db	1048	TACAAAGTACTC 1058	

RESULT 12  
US-10-425-114-10772  
; Sequence 10772, Application US/10425114  
; Publication No. US20040034888A1

```

: GENERAL INFORMATION:
: APPLICANT: Liu, Jingdong
: APPLICANT: Zhou, Yihua
: APPLICANT: Kovacic, David K.
: APPLICANT: Screen, Steven E.
: APPLICANT: Tabaska, Jack E
: APPLICANT: Cao, Yongwei
: TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
: TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
: FILE REFERENCE: 38-21(53313)B
: CURRENT APPLICATION NUMBER: US/10/425,114
: CURRENT FILING DATE: 2003-04-28
: NUMBER OF SEQ ID NOS: 73128
: SEQ ID NO 10772
: LENGTH: 1145
: TYPE: DNA
: ORGANISM: Glycine max
: FEATURE:
: OTHER INFORMATION: Clone ID: 700970311_FLI
: US-10-425-114-10772

```

Query Match	35.6%	Score 450.4	DB 18	Length 1145
Best Local Similarity	67.8%	Pred. NO. 6.6e-117		
Matches 660	Conservative	0	Mismatches 311	Indels 3
				Gaps 2

Qy	1	ACGCTAAAGTAAAGAAAAATGGGACGCTTAAAGTTTCTGATTCTGCTCGCTCTT	73
Db	34	AAACAAAAACAGAACACATATGGACACTTAAAGTTCTCAGCTCTTCCCCCGTTG	93
Qy	74	CTGATGATGCTGAGCAATTGAGAACCGCTTTTGAAGATGGGGTACGAACGAGACTTGA	133
Db	94	CAGATGATTTGTGAGCAGCTCCGAAAAAGCTTCTCAGTTGGGGAACTTAACGAGAGCTGA	153
Qy	134	TCAATATCAATCTTGGGCTCAGAAAGTCGAAACAGAGAAAGTATCATCGGGAACATACC	193
Db	154	TTGATATGATCTTGGCTCATAGGAATGCTGTCAAGAGAACTTTCAGAGAACTTAATG	213
Qy	194	ACGAAACCTTACGGCGGAGACCTTCTCAAGACTCTTGACAAGAGCTCTTAACATTTTCG	253
Db	214	CCGACAGCTTATGGAGAAAGATCTCTCAAGGCTTGGACAAAGAACTCAGAGATATTTTG	273
Qy	254	AGAGAGCTATCTTTGTTGGGACTCTTGAACCCGGTGAAGCGTGATGCTTATTTGGCTAATG	313
Db	274	AGAGGCTGTGTTTCATCTTTTGAGACCTTGATTTCTGTGAACGTGATGCAATTTTGGCAACG	333
Qy	314	AAAGCTACAAAAGATAGTAGCTTCAAGCAACCAAGTTCTTATGGAAAGTTGCTTGACAAGGA	373
Db	334	AAAGGACTTAAATATGACCTTCAAGCAATCAGGTTCTGGTGAATATACCTGCACTAAGT	393
Qy	374	CATCAACGCGAGCTGCTTCAAGCTTAGGCAAGCTTACATGCTGCTACAGAAAGTCTCTTG	433
Db	394	CCCTGTGAACAATGTTTGGCTCGAGGAAAGGCTTACATGTTCTTATTAAGAAAGCTCTGG	453
Qy	434	AAAGAGCGTTGCTGCACAACCTACCGGAGATTTGAGAAAGCTTTTGGTTCTCTGTGA	493
Db	454	AGGAGGATGTTGCTCATCAACAACGAGGACTTCCGTAAGCTCATCACTCTGTGGTTA	513
Qy	494	CCTCATACAGTTCGAGAGGAGATGAAAGTGAACATGACATTTGGCTPAAGCAAGAACTAAGC	553
Db	514	GTTCTTATGATATGAAAGGAGATGAGGTCAATTGACCT-6CAAAAACGTGACAAAAT	572
Qy	554	TGCTCATGAGAAATCAAGGACAGGACTACATGATGAGAGTGTATTGAACTTGT	613
Db	573	TGCTGATGAGAAATTTCAACAGGCTTATTAAGATGAAGACTTCACTCAGGATTTTGG	632
Qy	614	CCACAAGAAAGCAAGCTCAGATCAATGTCAATTTTAAACGTTACCAAGATATCATGGCG	673
Db	633	CCACAAGGACAGAGACAGATTAATGTAATCTTTGATTCACATACAAAGATGATTTGG	692
Qy	674	AGGAATTTCTCAAGAGCTTTGAGAAAGAGATGATGATGACAAAGTTCTTGCACTTTGA	733
Db	693	ACAAGATATCAACAGAGACTGA--AGGCTGATCCAAAGAGAGAGTTCTGTGCTTTACTAA	750

[illegible]

RESULT 13  
US-10-424-599-70457  
; Sequence 70457, Application US/10424599

```

: GENERAL INFORMATION:
: APPLICANT: La Rosa Thomas J
: APPLICANT: Kovalic David K
: APPLICANT: Zhou Yihua
: APPLICANT: Cao Yongwei
: TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
: TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
: FILE REFERENCE: 38-21(53223)B
: CURRENT APPLICATION NUMBER: US/10/424,599
: NUMBER OF SEQ ID NOS: 285684
: SEQ ID NO 70457
: LENGTH: 1321
: TYPE: DNA
: ORGANISM: Glycine max
: FEATURE:
: OTHER INFORMATION: Clone ID: PAT_MRT3847_34633C.1
US-10-424-599-70457

Query Match      35.0%; Score 430.8; DB 18; Length 1321;
Best Local Similarity 66.9%; Pred. No. 2,7e-111;
Matches 644; Conservative 0; Mismatches 312; Indels 6; Gaps 2;

Qy      26  GAAGAAAAATGGGCACTCTTAAAGTTCGTATTCCTGCTCTCTTGTATGATGATCTG 85
Db      82  GATGAGAAAATGGCACTCTTAAAGTTCGTATTCCTGCACTCTTCTTCATTTGGAAGACGTG 141

Qy      86  AGCAATTGAGAACCGCTTTTGAAGATGGGCTACGAACGAGACTTGATCATATCAATCT 145
Db      142 AGCAATTGCAAAAAGCTTTTCAAGGGTGGGGAACGATGAAAGATTGATATATCAATCC 201

Qy      146 TGGCTCACAAGAGCTGGAACAGGAAATCATACGGCAAGCATTCACAGAAACCTACG 205
Db      202 TGGGTTCACAGAAATGGCGCTCAGGCTAAGTTGATGCTGAAGCTTATATGACAAACCATG 261

Qy      206 GCGAAGACTTCTTCAGACTCTTGACAAAGAGCTCTCTTAACGATTTGAGAGAGCTATCT 265
Db      262 GAGAAAGATCTTCTTAAAGATTTAGACAAAGAACTTTCAAGTGACTTTGACCGGGCTGTGT 321

Qy      266 TGTGTGGACTCTTGAACCCGGTGAAGCGTGAATGCTTATTGGCTAATGAAGTCAAAAA 325
Db      322 TGTGTGGAATCTTGATCTGCTGAAGGTATGCTTTTAAAGTAAATGAAGCCACTAAGA 381

Qy      326 GATGACTCTTCAAGCAACGATCTTATGGAAGTGTGTCACAAAGACATCAACGACG 385
Db      382 TGTGTGACTTCAACAACATTTGGGTCAATTTTGGAAATTTGCTTCCACTAGATCTTCACTTGATC 441

Qy      386 TGCTTCAAGCTTAAGCAAGCTTACCATGCTGCTACAAAGATGCTTTGAAGAGACGTGG 445

```

```

Db      442  TCCTTAAGGCAGAACAGCATATCAAGCCCGTTTCAAAAAGTCCCTTGAAGAGACGTTG  501
OY      446  CTCACCACTATACCGGTGACTTCAGAAAAGCTTTTGGTTTCTCTTGTATCTATACAGGT  505
Db      502  CCTATCATCTATAAAGGGGACATCCGAAAGCTATTAAGTTCCTCTTGTGACATATTCGTT  561
OY      506  ACGAAGAGAGA-TGAAGTGAACATGACATTTGGCTTAGCAAGAAAGCTAAGTGGTCCATGAG  564
Db      562  ATGAGAGGGGATTTGAGGTGAACATGACATTTGGCAAAAATCTGAGGCTTAAACTGTCTTAC  621
OY      565  AAAATCAAGSACAAAGCACTACATGATGATGAGATGTTATAGAACTTGTGCCAAGAAAG  624
Db      622  AAGATTGCAGAGAAAGGCCCTACATGATGATGAGAGCTGATCAGGATTTTGTCAACAGAGT  681
OY      625  AAAAGTCAGATCAATGCTACTTTTAAACCGTTAACAGATGATCANGCGAGGAAATTC  684
Db      682  AAAGCAGAGTTAAAGCAACCACTAAATCAGTACATCAATGATGGGGAATGCCATTAAC  741
OY      685  AAGAGCTCTTGAGGAAGAGATGATGATGACAAAGTTCCTTGCACTTTGA-----GGTCAA  739
Db      742  AAGGATCTGAAAAAATTTGAATCTCAAAAAGATGTGAATATCTGCAATTTTGAAGGGCAAG  801
OY      740  CCATTACAGTGTCTTGACAAGACAGAGCTTTACTTGTGATGTTCTTCCGTTGAGCAATCA  799
Db      802  CCATTAAAGTGTCTTGACTTACCTCGAGAAATATTTTGCAAAAAGCTTGAAGGTGGCTATA  861
OY      800  ACAAAAATCTGAACTGATGAGAGAGCACTCAGTAAATTTGTGACACAGAGCTGAGATTG  859
Db      862  ACAAGCTGAGAACTGATGAGAGGGGCCCTTACTAGATGATGTGACAAACAGAGCTGAGTTG  921
OY      860  ACTTGAAGGTCAATTGAGAGAGAGTACCAAGCGCAGAAACAGCATTCCTTTGAGAAAGCTA  919
Db      922  ATTTCAGAGGAATTTGCTGAGGAATACCAAAAGAAAGAAACAGCATTCCTCTTGAGACGG  981
OY      920  TTACCAAGAGACACTCGTGGAGATTACGAGAAAGTCTCGTCCACTTCTCGTGAAGATG  979
Db      982  TAGCCAGTGAACACTTCTGAGAGCTATCAGAGGATTCCTCTTGCACTGTTGAGCATGAGG  1041
OY      980  AT 981
Db      1042  AT 1043

RESULT 14
US-10-219-220-20
; Sequence 20, Application US/10219220
; Publication No. US20030082724A1
; GENERAL INFORMATION:
; APPLICANT: Flinn, Barry
; APPLICANT: Lasham, Annette
; TITLE OF INVENTION: Compositions affecting programmed cell
; TITLE OF INVENTION: death and their use in the modification of plant development
; FILE REFERENCE: 11000.1022c1
; CURRENT APPLICATION NUMBER: US/10/219, 220
; CURRENT FILING DATE: 2002-08-14
; PRIOR APPLICATION NUMBER: U.S. No. US20030082724A1 09/325, 932
; PRIOR FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 290
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 20
; LENGTH: 789
; TYPE: DNA
; ORGANISM: Pinus radiata
US-10-219-220-20

```

Query Match 29.6%; Score 364.2; DB 14; Length 789;  
Best Local Similarity 69.7%; P<sub>rod</sub> 1.7e-92;  
Matches 508; Conservative 0; Mismatches 218; Indels 3; Gaps 1

OY	85	GACCAATTGAAACCGCTTTTAAAGGATGGGGTACGAAACGAGACTTGATCATATCAATC	144
Db	124	GACACAGCTCCAAAACCTTTCCGACAGTGGGGGAGAAATGAATCTTGATCATATTCATA	183
OY	145	TTGGCTACAGAAAGTGTCTGAAACAGAGAAAGTCATACGACCAAGCATACACGAAACCTAC	204
Db	184	CTGCTCTCACGAAACGACGACGAGGAAAGTAATCCGACAAATATGCGGAGACATAT	243
OY	205	GGGGAAGACCTTCTCAAGCTTTTGAACAAGAGCTCTCTAACATTTTCGAGAGCTATC	264
Db	244	GGGAAAGATCTTCTCAAAAGCGTTTACAGAGAACTCTTAAGTACTTTGAGAGATCTGTG	303
OY	265	TTGTGTGAGACTTTGAACCCGCTAGCGTGATGCTTATTTGGCTAATGAAGCTACAAA	324
Db	304	CTTCTGTGAGACCTTGATCTCTGCGAGGCTATGTGATTTCTTGTCCAAATGAAGCTACCAAG	363
OY	325	AGATGACCTTCAAGCAACCAAGTTCTTATGGAAGTTGCTTGCAACAAGACATCAAGCAG	384
Db	364	AGATTGACTTCGAGCAACTGGGTTCTCATGAAATTTGCTTCAAGAGGCTTTCAMTGGAG	423
OY	385	CTGCTTCAGCTAGGAGCAAGCTTACCATGCTGCGCTCAAGAAGTCTCTTGAAGAGAGCTT	444
Db	424	TTATTATGTGTGAGGAGAGGCTTATATGCTGCTTATGAAATCTCTTGAAGAAGACATC	483
OY	445	GCTCACACACATACCGGTACTTCAGAAAGCTTTTGGTTTCTCTGTTACTATACAGG	504
Db	484	GCATATCACACTACTGGGATTTCCGCAAGCTGCTGTTCTCTGCGAAGTACTTTCGG	543
OY	505	TACAGAGAGATGAAATGAAACATGACATTTGGCTAAGCAAGAAGCTAAGCTGTGTCATGAG	564
Db	544	TATAGAGGGCTGAGGTGAACATGACATTTGGGAGATAGAGGCTTAAGATCTTCATGAG	603
OY	565	AAATCAAGGACACGACTTACATGATGATGAGATGTTATTAGAACTTGTCCACAAAGAC	624
Db	604	AAGTTCACAGAAAGGCTTACATCATGATGAGCTCATCAAAATTTGTACTACAAGAAAT	663
OY	625	AAAGCTCAATCAATGCTTCTTTAAACGTTTACCAAGATGATCANTGGCGAGAAATCTC	684
Db	664	AAAGCTCAGTTATATCAACCTCAATTACTTCAACAATGAGTTTGGAAATGCCATAC	723
OY	685	AAAGTCTTGAGGAGAGAGATGATGATGACAAAGTTCTTGCACCTTTTGAAGTCAACATT	744
Db	724	AAAGATCT---GAAAGCTGATCCAAATGTAATTTTGAAACCTGCTGAGATCAGCAATT	783
OY	745	CAGTGTCTG	753
Db	781	AAAGTCTTG	789

```

RESULT 15
US-10-393-840-26
/ Sequence 26, Application US/10393840
/ Publication No. US20030229922A1
/ GENERAL INFORMATION:
/ APPLICANT: Bloksberg, Leonard N.
/ TITLE OF INVENTION: Materials and Methods for the
/ TITLE OF INVENTION: Modification of Plant Cell Wall Polysaccharides
/ FILE REFERENCE: 11000.1012c3
/ CURRENT APPLICATION NUMBER: US/10/393,840
/ CURRENT FILING DATE: 2003-03-20
/ PRIOR APPLICATION NUMBER: US 09/636,800
/ PRIOR FILING DATE: 2000-08-10
/ PRIOR APPLICATION NUMBER: US 09/170,862
/ PRIOR FILING DATE: 1998-10-13
/ PRIOR APPLICATION NUMBER: US 60/148,426
/ PRIOR FILING DATE: 1999-08-11
/ PRIOR APPLICATION NUMBER: PCT NZ/99/00169
/ PRIOR FILING DATE: 1999-10-08
/ NUMBER OF SEQ ID NOS: 956
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 26
/
/ LENGTH: 789

```

TYPE: DNA  
ORGANISM: Eucalyptus grandis  
US-10-393-840-26

Query Match 29.6%; Score 364.2; DB 17; Length 789;  
Best Local Similarity 69.7%; Pred. No. 1.7e-92;  
Matches 508; Conservative 0; Mismatches 218; Indels 3; Gaps 1;

```
QY 25 AGAAGAAAATGGGACTCTTAAGGTTTCTGATTCCTGCTCTCTGATGATGCT 84
   |||
Db 64 AGAAGAGAAATGGGACTATCGCGATGCGACCTCGGTCCGCTCCGCTGAGATGCC 123
   |||
QY 85 GAGCAATTGAGACCGCTTTGAAGGATGGGGTACGACGAGGACTGATCATATCATC 144
   |||
Db 124 GAGCAGCTCCAAAAGCTTTGACAGATGGGGAGCAATGAGATCTGATCATATCCATA 183
   |||
QY 145 TTGGCTCAGAGATGCTGAAACAGAGAAATCATCAGCAAGCATACAGAACTTAC 204
   |||
Db 184 CTGCTCAGAGAAAGGCGGCGGCAAGTAATCCGACAAACATATGCCGAGACATAT 243
   |||
QY 205 GGGCAAGACTTCTCAAGACTTTGACAGAGACTCTCTAACGATTTGAGAGACTATC 264
   |||
Db 244 GGGGAGATCTTCTCAAGCGCTTGACAGGAACTCTTAGTGACTTGAAGATCTGTG 303
   |||
QY 265 TTGTTGGACTCTTGAACCGGATGAGGCTGATGCTTATGCTAATGAAAGCTACAAA 324
   |||
Db 304 CTTCTGTGAGCCTGAGATCTGCGAGGATGATCATCTTGTCAAATGAAAGCTACAG 363
   |||
QY 325 AGATGACTTCAGAACCAAGTTCTTAAGAAAGTTGCTTGACAGAGACATCAACGAG 384
   |||
Db 364 AGATTGACTTGAGCAACTGGGTTCTCATGSAATGCTTGACAGAGCTCTTCAATGAG 423
   |||
QY 385 CTGCTTACGCTGAGCAAGCTTACCATGCTGCTACAAAGATCTTTGAAGAGACGTT 444
   |||
Db 424 TTATTCAATGATGAGGACAGCCTATCATGCTGTTAATGAATCTTTGAAGAACATC 483
   |||
QY 445 GCTTACCACTACCGGATGATTCAGAAAGCTTTGGTTCTCTTTACTCATACAG 504
   |||
Db 484 GCATATCACACTACTGGGATTTCCGCAAGCTGCTGTGTCTTGCAAGTACCTTTGG 543
   |||
QY 505 TACGAGGAGATGAGTGAATGATGATGATGCTTAAGCAAGAGCTAAGCTGTCATGAG 564
   |||
Db 544 TATGAGGGGCTGAGTGAATGATGATGATGCTTAAGCAAGAGCTAAGATCTTCAATGAG 603
   |||
QY 565 AAAATCAAGACAAAGCACTACATGATGAGATGTTAATGAAATCTTGTCCAAAGAGC 624
   |||
Db 604 AAGATTACGAGAGAGGCTTACATCATGATGAGCTCATCAGAAATTTTACTACAAAGAT 663
   |||
QY 625 AAAGCTCAGATCATGCTACTTTTAACCGTTACCAAGATGATCATGCGAGGAAATCTC 684
   |||
Db 664 AAAGCTCAGCTTAATGCAACCTCAATTAACCAAGATGATGTTGGAAATGCCATCAAC 723
   |||
QY 685 AAGAGCTTGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 744
   |||
Db 724 AAGGATCT---GAAGGCTGATCCAAATGATGATTTTCTGAAGACTGCTGATGATGAT 780
   |||
QY 745 CAGTCTTG 753
   |||
Db 781 AAGTCTTG 789
   |||
```

Search completed: August 22, 2005, 07:12:44  
Job time : 2743 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 19, 2005, 20:23:59 ; Search time 165 Seconds

(without alignments)  
743.049 Million cell updates/sec

Title: US-10-690-564-2

Perfect score: 1611  
Sequence: 1 MATLKVSDEVPAIPSDDAEQL.....KOTRGDYKMLVALGEDIA 317

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

A\_Geneseq\_16Dec04:\*  
1: geneseq1980s:\*  
2: geneseq1990s:\*  
3: geneseq2000s:\*  
4: geneseq2001s:\*  
5: geneseq2002s:\*  
6: geneseq2003as:\*  
7: geneseq2003bs:\*  
8: geneseq2004s:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1611	100.0	317	3	AAG09200	Aag09200 Arabidops
2	1611	100.0	317	3	AAG08874	Aag08874 Arabidops
3	1611	100.0	317	8	ADU63872	Adj63872 Plant 1lp
4	1611	100.0	317	8	ADN74567	Adn74567 Thale cre
5	1611	100.0	327	3	AAG09199	Aag09199 Arabidops
6	1611	100.0	328	3	AAG08873	Aag08873 Arabidops
7	1185.5	73.6	316	3	AAB16294	Aab16294 Eucalyptu
8	1171.5	72.7	316	3	AAB16321	Aab16321 Pinus rad
9	1171.5	72.7	316	4	AAB55730	Aab55730 Annexin-1
10	1171.5	72.7	316	7	ADB94698	ADB94698 Programme
11	1132.5	70.3	316	7	ADB94896	ADB94896 Programme
12	1071	66.5	211	3	AAG09201	Aag09201 Arabidops
13	1071	66.5	211	3	AAG08875	Aag08875 Arabidops
14	1071	66.5	211	3	AAG06311	Aag06311 Arabidops
15	1067	66.2	317	3	AAG33493	Aag33493 Arabidops
16	1067	66.2	317	3	AAG49202	Aag49202 Arabidops
17	1067	66.2	317	8	ADN74791	Adn74791 Thale cre
18	1034.5	64.2	316	3	AAG37921	Aag37921 Arabidops
19	988	61.3	302	3	AAG49231	Aag49231 Arabidops
20	988	61.3	304	3	AAG49230	Aag49230 Arabidops
21	928.5	57.6	284	3	AAG37922	Aag37922 Arabidops
22	866.5	53.8	239	3	AAB16292	Aab16292 Eucalyptu
23	866.5	53.8	239	4	AAB16324	Aab16324 Pinus rad
24	866.5	53.8	239	4	AAB55733	Aab55733 Annexin-1
25	866.5	53.8	239	7	ADB94701	ADB94701 Programme

#### ALIGNMENTS

26	787	48.9	240	3	AAG49232	Aag49232 Arabidops
27	787	48.9	240	3	AAG33494	Aag33494 Arabidops
28	787	48.9	240	3	AAG49203	Aag49203 Arabidops
29	784.5	48.7	239	3	AAG37923	Aag37923 Arabidops
30	770	47.8	153	3	AAG06312	Aag06312 Arabidops
31	721	44.8	242	3	AAB16293	Aab16293 Pinus rad
32	712	44.2	220	3	AAG33495	Aag33495 Arabidops
33	712	44.2	220	3	AAG49204	Aag49204 Arabidops
34	625	38.8	204	3	AAG17531	Aag17531 Arabidops
35	625	38.8	206	3	AAG17530	Aag17530 Arabidops
36	602.5	37.4	321	3	AAG30142	Aag30142 Arabidops
37	602.5	37.4	321	3	AAG38352	Aag38352 Arabidops
38	602.5	37.4	345	3	AAG38351	Aag38351 Arabidops
39	573	35.6	184	4	AAB16325	Aab16325 Pinus rad
40	573	35.6	184	4	AAB55734	Aab55734 Annexin-1
41	573	35.6	184	7	ADB94702	ADB94702 Programme
42	563	34.9	184	3	AAB16290	Aab16290 Pinus rad
43	512	31.8	316	7	ADJ69605	Adj69605 Human hea
44	509	31.6	327	4	ABBI0307	Abbi0307 Human CDV
45	509	31.6	327	5	ABP66894	Abp66894 Human pol

ALIGNMENTS

RESULT 1	
AA09200	standard; protein; 317 AA.
XX	AA09200;
XX	AA092200;
XX	
DT	17-OCT-2000 (first entry)
XX	
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 7039.
XX	
KW	Protein identification; signal transduction pathway; metabolic pathway;
KM	hybridisation assay; genetic mapping; gene expression control; promoter;
KW	termination sequence.
XX	
OS	Arabidopsis thaliana.
XX	
PN	EP1033405-A2.
XX	
PD	06-SEP-2000.
XX	
PP	25-FEB-2000; 2000EP-00301439.
XX	
PR	25-FEB-1999; 99US-0121825P.
PR	05-MAR-1999; 99US-0123180P.
PR	09-MAR-1999; 99US-0123548P.
PR	23-MAR-1999; 99US-0125788P.
PR	25-MAR-1999; 99US-0126264P.
PR	29-MAR-1999; 99US-0126785P.
PR	01-APR-1999; 99US-0127462P.
PR	06-APR-1999; 99US-0128234P.
PR	08-APR-1999; 99US-0128714P.
PR	16-APR-1999; 99US-0129845P.
PR	19-APR-1999; 99US-0130077P.
PR	21-APR-1999; 99US-0130449P.
PR	23-APR-1999; 99US-0130510P.
PR	28-APR-1999; 99US-0130891P.
PR	30-APR-1999; 99US-0131449P.
PR	30-APR-1999; 99US-0132048P.
PR	30-APR-1999; 99US-0132407P.
PR	04-MAY-1999; 99US-0132484P.
PR	05-MAY-1999; 99US-0132485P.
PR	06-MAY-1999; 99US-0132486P.
PR	07-MAY-1999; 99US-0132487P.
PR	11-MAY-1999; 99US-0132563P.
PR	14-MAY-1999; 99US-0134218P.
PR	14-MAY-1999; 99US-0134219P.
PR	14-MAY-1999; 99US-0134221P.



PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149366P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153707P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.

PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161922P.  
PR 28-OCT-1999; 99US-0161993P.  
PR 29-OCT-1999; 99US-0162142P.  
  
Query Match 100.0%; Score 1611; DB 3; Length 317;  
Best Local Similarity 100.0%; Pred. No. 1,2e-139;  
Matches 317; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 MATLKVDSPVAPSDDAEQRTAFEGWGTEEDLLISILARSAPQKRVIRQAHYETGED 60  
DB 1 MATLKVDSPVAPSDDAEQRTAFEGWGTEEDLLISILARSAPQKRVIRQAHYETGED 60  
QY 61 LKLTLDKELSDNDFRALILMTLBERDALLANEAATKRTSSNOVLMEVACTRTSTOLH 120  
DB 61 LKLTLDKELSDNDFRALILMTLBERDALLANEAATKRTSSNOVLMEVACTRTSTOLH 120  
QY 121 APQAYHARYKKSLEEDVAHHTTGFRLKLVSLVTSYREGEDEVNMTLAKQBAKLVBKIK 180  
DB 121 APQAYHARYKKSLEEDVAHHTTGFRLKLVSLVTSYREGEDEVNMTLAKQBAKLVBKIK 180  
QY 181 DKXINDVDVIRILSTRKAQINATFNRYODDHGEELIKSLBEGDDDKFLALLRSTIOCL 240  
DB 181 DKXINDVDVIRILSTRKAQINATFNRYODDHGEELIKSLBEGDDDKFLALLRSTIOCL 240  
QY 241 TRPELYEVVDLRSAINITGTDEGALTRIVTTRABRIDLKVIGEYQRRNSIPLKKAIRKDT 300  
DB 241 TRPELYEVVDLRSAINITGTDEGALTRIVTTRABRIDLKVIGEYQRRNSIPLKKAIRKDT 300  
QY 301 RGDYKMLVALLGEDDA 317  
DB 301 RGDYKMLVALLGEDDA 317  
  
RESULT 2  
AAG08874  
ID AAG08874 standard; protein; 317 AA.  
XX  
AC AAG08874;  
XX  
DT 17-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 6582.  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-00301439.  
XX  
PR 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.

PR 23-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135124P.  
PR 24-MAY-1999; 99US-0135533P.  
PR 25-MAY-1999; 99US-0135629P.  
PR 27-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.

Query Match	Similarity	100.0%	Score	1611	DB 3	Length	317
Best Local	Similarity	100.0%	Pred. No.	1.2e-139			
Matches	317	Conservative	0	Mismatches	0	Indels	0
						Gaps	0
QY	1	MATLKVSQSDVPAPSDDAEQLRTAFEGWGCTNEDLIISLAHRSABQKRVIRQAYHETYGED	60				
DB	1	MATLKVSQSDVPAPSDDAEQLRTAFEGWGCTNEDLIISLAHRSABQKRVIRQAYHETYGED	60				
QY	61	LKLTLDKEUSNDPERAILMTLEPGERDALLANETKRWISSNOYLMEVACTRSTOLLH	120				
DB	61	LKLTLDKEUSNDPERAILMTLEPGERDALLANETKRWISSNOYLMEVACTRSTOLLH	120				
QY	121	ARQAYHARKYSLEEVVAHHTTGDRPKLVLSVTSYRGEQEVNMTLAKQBAKLVHEKIK	180				
DB	121	ARQAYHARKYSLEEVVAHHTTGDRPKLVLSVTSYRGEQEVNMTLAKQBAKLVHEKIK	180				
QY	181	DKHYNDEDVIRILSTRSKAQINATFNRYODDHGEIILKSLEEGDDDDKFLALNSTIOCL	240				
DB	181	DKHYNDEDVIRILSTRSKAQINATFNRYODDHGEIILKSLEEGDDDDKFLALNSTIOCL	240				
QY	241	TRPELYFVDVLRSAIKNTGTDEGALTRIVTTAEITDLKVIGEYQRRNSIPILEKATIKDT	300				
DB	241	TRPELYFVDVLRSAIKNTGTDEGALTRIVTTAEITDLKVIGEYQRRNSIPILEKATIKDT	300				
QY	301	RGDYKXKLVALLGEEDA	317				
DB	301	RGDYKXKLVALLGEEDA	317				
RESULT 3							
ADJ63872	ADJ63872	standard	protein	317	AA		
AC	ADJ63872						
XX							
DT	06-MAY-2004	(first entry)					
XX							
DE	Plant lipid metabolism protein OO-10 SEQ ID NO:74.						
XX							
KW	plant; lipid metabolism protein; LMP; seed storage compound;						
KW	transgenic plant.						
XX							
OS	Unidentified.						
XX							
PN	MO2004013304-A2.						
XX							
PD	12-FEB-2004.						
XX							
PF	04-AUG-2003; 2003MO-US024364.						
XX							
PR	02-AUG-2002; 2002US-040803P.						

XX (BADI ) BASF PLANT SCI GMBH.  
 PA Miltendorf V, Haertel HA, Bauer J, Oswald O;  
 PI WPI; 2004-157121/15.  
 DR N-PSDB; ADJ63871.  
 XX  
 PT New lipid metabolism proteins and nucleic acids, useful in producing  
 PT transgenic plants with increased levels of seed storage compound, e.g.  
 PT lipid, a fatty acid, a starch or a seed storage protein.  
 PS Claim 1; SEQ ID NO 74; 115bp; English.  
 XX  
 CC The invention relates to novel isolated lipid metabolism proteins (LMP)  
 CC and encoding nucleic acids comprising a polynucleotide sequence encoding  
 CC a polypeptide that functions as a modulator of seed storage compounds in  
 CC a plant. The LMP nucleic acid is useful in producing transgenic plants  
 CC with increased levels of seed storage compound, e.g. lipid, a fatty acid,  
 CC a starch or a seed storage protein, as markers for specific regions of  
 CC the genome and for evolutionary and protein structural studies. The  
 CC present sequence represents an LMP of the invention.  
 CC  
 XX Sequence 317 AA;  
 SQ  
 Query Match 100.0%; Score 1611; DB 8; Length 317;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-139;  
 Matches 317; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATLKVSQSVAPSDAEOQLRTAFEGWGTNEDLIISILAHRSAEQRKVIROAHYHETGED 60  
 DB 1 MATLKVSQSVAPSDAEOQLRTAFEGWGTNEDLIISILAHRSAEQRKVIROAHYHETGED 60  
 QY 61 LKTLDELSNDPERAILMTLPEGERDALIANEATKRWTSNQVLMVACTRSTQLH 120  
 DB 61 LKTLDELSNDPERAILMTLPEGERDALIANEATKRWTSNQVLMVACTRSTQLH 120  
 QY 121 AROAHARVKKSLSEEDVAHHHTGDFRKLVSIVTSYRYEGDEVNMTLAKOEKLVHEKIK 180  
 DB 121 AROAHARVKKSLSEEDVAHHHTGDFRKLVSIVTSYRYEGDEVNMTLAKOEKLVHEKIK 180  
 QY 181 DKHYNDEDEVIRIILSTRSKAQINATFNRYODHGEIILSLSEGGDDDKFALLRSTIQL 240  
 DB 181 DKHYNDEDEVIRIILSTRSKAQINATFNRYODHGEIILSLSEGGDDDKFALLRSTIQL 240  
 QY 241 TRPELYFVDVLRSAINKTGDEGALTRIVTTTBAIDLVKIGEEYORRNSIPLEKAITKOT 300  
 DB 241 TRPELYFVDVLRSAINKTGDEGALTRIVTTTBAIDLVKIGEEYORRNSIPLEKAITKOT 300  
 QY 301 RGDYERKMLVALLGEDDA 317  
 DB 301 RGDYERKMLVALLGEDDA 317

RESULT 4  
 ADN74567  
 ID ADN74567 standard; protein; 317 AA.  
 XX  
 AC ADN74567;  
 DT 15-JUL-2004 (first entry)  
 XX  
 DE Thale cress protein repressed in E2Fa/Dpa expressing plants SeqID 2462.  
 XX  
 KW plant; transgenic; E2Fa/Dpa transcription factor; growth regulator;  
 KW animal feed product; thale cress; cell wall biosynthesis;  
 KW nitrogen metabolism; carbon metabolism.  
 XX  
 OS Arabidopsis thaliana.  
 XX  
 XX WO2004035798-A2.  
 XX  
 PD 29-APR-2004.

XX  
 PF 20-OCT-2003; 2003WO-EP011658.  
 XX  
 PR 18-OCT-2002; 2002EP-00079408.  
 XX  
 PA (CROP-) CROPPDESIGN NV.  
 XX  
 PI Inze D, De Veylder L, Vlieghe K;  
 XX  
 DR WPI; 2004-348466/32.  
 DR N-PSDB; ADN74566.  
 XX  
 PT Altering plant characteristics, useful for producing plants for enzyme or  
 PT pharmaceutical production comprises modifying in a plant, expression of  
 PT one or more nucleic acids and/or modifying level or activity of one or  
 PT more proteins.  
 PS Claim 1; SEQ ID NO 2462; 134bp; English.  
 XX  
 CC This invention relates to a novel method for altering one or more plant  
 CC characteristics. Specifically, it refers to identifying genes that are up  
 CC - or down-regulated in transgenic plants overexpressing the heterodimeric  
 CC E2Fa/Dpa transcription factor of Arabidopsis and using these sequences to  
 CC alter plant characteristics accordingly. The present invention describes  
 CC generating transgenic plants for the production of growth regulators,  
 CC enzymes, therapeutics, pharmaceuticals and animal feed products, where  
 CC the altered plant characteristics are selected from increased yield or  
 CC biomass, enhanced survival capacity, stress tolerance, plant architecture  
 CC or physiology, altered lipid mobilization, biochemistry, signal  
 CC transduction, storage lipid mobilization and/or altered photosynthesis,  
 CC each relative to the corresponding wild type plants. Accordingly, these  
 CC sequences can also be useful as positive or negative selectable markers  
 CC during transformation of cells or tissues. The identified genes play a  
 CC role in a variety of biological processes such as DNA replication, cell  
 CC wall biosynthesis, nitrogen and/or carbon metabolism or they function as  
 CC transcription factors. This polypeptide sequence is thale cress protein  
 CC expressed by a gene repressed 1.3 fold or more in plants overexpressing  
 CC the E2Fa/Dpa transcription factor, given in an exemplification of the  
 CC invention.  
 CC  
 XX Sequence 317 AA;  
 SQ  
 Query Match 100.0%; Score 1611; DB 8; Length 317;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-139;  
 Matches 317; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATLKVSQSVAPSDAEOQLRTAFEGWGTNEDLIISILAHRSAEQRKVIROAHYHETGED 60  
 DB 1 MATLKVSQSVAPSDAEOQLRTAFEGWGTNEDLIISILAHRSAEQRKVIROAHYHETGED 60  
 QY 61 LKTLDELSNDPERAILMTLPEGERDALIANEATKRWTSNQVLMVACTRSTQLH 120  
 DB 61 LKTLDELSNDPERAILMTLPEGERDALIANEATKRWTSNQVLMVACTRSTQLH 120  
 QY 121 AROAHARVKKSLSEEDVAHHHTGDFRKLVSIVTSYRYEGDEVNMTLAKOEKLVHEKIK 180  
 DB 121 AROAHARVKKSLSEEDVAHHHTGDFRKLVSIVTSYRYEGDEVNMTLAKOEKLVHEKIK 180  
 QY 181 DKHYNDEDEVIRIILSTRSKAQINATFNRYODHGEIILSLSEGGDDDKFALLRSTIQL 240  
 DB 181 DKHYNDEDEVIRIILSTRSKAQINATFNRYODHGEIILSLSEGGDDDKFALLRSTIQL 240  
 QY 241 TRPELYFVDVLRSAINKTGDEGALTRIVTTTBAIDLVKIGEEYORRNSIPLEKAITKOT 300  
 DB 241 TRPELYFVDVLRSAINKTGDEGALTRIVTTTBAIDLVKIGEEYORRNSIPLEKAITKOT 300  
 QY 301 RGDYERKMLVALLGEDDA 317  
 DB 301 RGDYERKMLVALLGEDDA 317

RESULT 5  
 AAG09199

ID AAG09199 standard; protein; 327 AA.  
XX  
AC AAG09199;  
XX  
DT 17-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 7038.  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-00301439.  
XX  
PR 25-FEB-1999; 99US-0121625P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 28-APR-1999; 99US-0130891P.  
PR 30-APR-1999; 99US-0131449P.  
PR 04-MAY-1999; 99US-0132407P.  
PR 05-MAY-1999; 99US-0132484P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.

PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 24-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144633P.  
PR 20-JUL-1999; 99US-0144684P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147933P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148131P.  
PR 12-AUG-1999; 99US-0148346P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.

```
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 30-AUG-1999; 99US-0151080P.
PR 31-AUG-1999; 99US-0151303P.
PR 01-SEP-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159283P.
PR 13-OCT-1999; 99US-0159284P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161202P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.
```

Query Match 1 100.0%; Score 1611; DB 3; Length 327;  
Best Local Similarity 100.0%; Pred. No. 1,3e-139;  
Matches 317; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MATLKXSDSVPAQSDAEQRTAFEGWGTNEDLIISILHRSAEORRVIRQAHETTYGDD 60
DB 11 MATLKXSDSVPAQSDAEQRTAFEGWGTNEDLIISILHRSAEORRVIRQAHETTYGDD 70
QY 61 LKLTLDKELSNDFERRAILLWTLLEPGERDALLANEATKRWTSNQVIMEVACTRTSTQLH 120
DB 71 LKLTLDKELSNDFERRAILLWTLLEPGERDALLANEATKRWTSNQVIMEVACTRTSTQLH 130
QY 121 ARQAVYARVYKKSLEEDVAHHTTDFRKLVSIVTSYRYEDEVNMTLAEQALVHEKIK 180
DB 131 ARQAVYARVYKKSLEEDVAHHTTDFRKLVSIVTSYRYEDEVNMTLAEQALVHEKIK 190
QY 181 DKHYNDEVIRILSTRSKAQINATFNRYODDGEELIKSLSEGGDDDKFLALLRSTIOCL 240
```

```
DB 191 DKHYNDEVIRILSTRSKAQINATFNRYODDGEELIKSLSEGGDDDKFLALLRSTIOCL 250
QY 241 TRPELYFVUVLSAINKTGTDBGALTRIVTTTAEIDLKVIGEBYORRNSIPLEKAITKOT 300
DB 251 TRPELYFVUVLSAINKTGTDBGALTRIVTTTAEIDLKVIGEBYORRNSIPLEKAITKOT 310
QY 301 RGDYKMLVALLGEDDA 317
DB 311 RGDYKMLVALLGEDDA 327

RESULT 6
AAG08873
ID AAG08873 standard; protein; 328 AA.
AC AAG08873;
XX
XX 17-OCT-2000 (first entry)
DT
DE Arabidopsis thaliana protein fragment SEQ ID NO: 6581.
KW Protein identification; signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
OS Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-00301439.
XX
XX 25-FEB-1999; 99US-0121825P.
XX 05-MAR-1999; 99US-0123180P.
XX 09-MAR-1999; 99US-0123548P.
XX 23-MAR-1999; 99US-0125788P.
XX 25-MAR-1999; 99US-0126264P.
XX 29-MAR-1999; 99US-0126785P.
XX 01-APR-1999; 99US-0127462P.
XX 06-APR-1999; 99US-0128234P.
XX 08-APR-1999; 99US-0128714P.
XX 16-APR-1999; 99US-0129845P.
XX 19-APR-1999; 99US-0130077P.
XX 21-APR-1999; 99US-0130449P.
XX 23-APR-1999; 99US-0130510P.
XX 23-APR-1999; 99US-0130891P.
XX 28-APR-1999; 99US-0131449P.
XX 30-APR-1999; 99US-0132048P.
XX 30-APR-1999; 99US-0132407P.
XX 04-MAY-1999; 99US-0132484P.
XX 05-MAY-1999; 99US-0132485P.
XX 06-MAY-1999; 99US-0132486P.
XX 06-MAY-1999; 99US-0132487P.
XX 07-MAY-1999; 99US-0132863P.
XX 11-MAY-1999; 99US-0132866P.
XX 14-MAY-1999; 99US-0134218P.
XX 14-MAY-1999; 99US-0134219P.
XX 14-MAY-1999; 99US-0134221P.
XX 18-MAY-1999; 99US-0134370P.
XX 19-MAY-1999; 99US-0134768P.
XX 19-MAY-1999; 99US-0134941P.
XX 20-MAY-1999; 99US-0135124P.
XX 21-MAY-1999; 99US-0135353P.
XX 24-MAY-1999; 99US-0135629P.
XX 25-MAY-1999; 99US-0136021P.
XX 27-MAY-1999; 99US-0136392P.
XX 28-MAY-1999; 99US-0136782P.
XX 01-JUN-1999; 99US-0137222P.
XX 03-JUN-1999; 99US-0137528P.
XX 04-JUN-1999; 99US-0137502P.
XX 07-JUN-1999; 99US-0137724P.
XX 08-JUN-1999; 99US-0138094P.
```

PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 15-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 28-JUL-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.

PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151067P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158363P.  
PR 13-OCT-1999; 99US-0158293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160982P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161922P.  
PR 28-OCT-1999; 99US-0161932P.  
PR 29-OCT-1999; 99US-0162142P.

Query Match 100.0%; Score 1611; DB 3; Length 328;  
Best Local Similarity 100.0%; Pred. No. 1.3e-139;  
Matches 317; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MATLKVSDVAPSDDAEQRLTAFFGMCNTNEDLIISILAHRSABQKVIROAHYHETGCD 60

```

Db      12 MATLKVSDSVAPASDDAEQLRTAFEGWGTMEDLIISILHRSAPQRKVIQAHYETGED 71
Qy      61 LKLTLDKELSNDFERRAILLWTLPEGERDALLANFATKRTWSSNOVLMEVACTRTSTQLH 120
Db      72 LKLTLDKELSNDFERRAILLWTLPEGERDALLANFATKRTWSSNOVLMEVACTRTSTQLH 131
Qy      121 AROAYHARYKKSLEEDVAHHTTGDFRKLIVSLVTSYRYEGDEVNMTLAKQBAKLVHEKIK 180
Db      132 AROAYHARYKKSLEEDVAHHTTGDFRKLIVSLVTSYRYEGDEVNMTLAKQBAKLVHEKIK 191
Qy      181 DKHYNDESVIRILSTRSKAQINATFNRYODHGEELIKSLSEGGDDDKFLALRSTIOCL 240
Db      192 DKHYNDESVIRILSTRSKAQINATFNRYODHGEELIKSLSEGGDDDKFLALRSTIOCL 251
Qy      241 TRPELYFVDVLRSAINKGTDEGALTRIVTTTAEIDLVKVIAGEYQRNSIPLEKAITKDT 300
Db      252 TRPELYFVDVLRSAINKGTDEGALTRIVTTTAEIDLVKVIAGEYQRNSIPLEKAITKDT 311
Qy      301 RGDYERKMLVALLGEDDA 317
Db      312 RGDYERKMLVALLGEDDA 328

RESULT 7
AAB16294
ID AAB16294 standard; protein; 316 AA.
AC AAB16294;
XX
XX 31-OCT-2000 (first entry)
XX
DE Eucalyptus grandis amexin protein sequence SEQ ID NO:56.
XX
XX Eucalyptus grandis; pinus radiata; Monterey pine; modification;
XX plant cell wall; polysaccharide; polysaccharide biosynthetic pathway;
XX transgenic plant.
XX
OS Eucalyptus grandis.
XX
XX WO200022092-A2.
XX
XX 20-APR-2000.
XX
XX 08-OCT-1999; 99WO-NZ000169.
XX
XX 13-OCT-1998; 98US-00170862.
XX
XX 11-AUG-1999; 99US-0148426P.
XX
XX (GENE-) GENESIS RES & DEV CORP LTD.
XX (FLET-) FLETCHER CHALLENGE FORESTS LTD.
XX
XX Bloksberg LN;
XX
XX WPI; 2000-339328/29.
XX
XX N-PSDB; AAB67100.
XX
XX New genes encoding proteins involved in a plant polysaccharide
XX biosynthetic pathway, useful for modulating or altering the
XX polysaccharide content, composition or structure of the plant.
XX
XX Claim 17, Page 66; 301pp; English.
XX
XX The present invention describes isolated polynucleotides (PN) comprising
XX a sequence selected from one of 835 nucleotide sequences given in
XX CC AAA67073 to AAA67907, their (reverse) complements, sequences producing an
XX CC Expectation (E) value of 0.01 or less compared to the 835 sequences
XX CC sequences at least 50% identical to them, 200, 100, 40 or 20-mers of the
XX CC 835 sequences or sequences that are degenerately equivalent or allelic to
XX CC the 835 sequences. The polynucleotides are used to modify the activity of
XX CC a polypeptide involved in a polysaccharide biosynthetic pathway in the
XX CC plant. They are especially used to modulate or alter the polysaccharide
XX CC content, composition or structure of the plant. AAB16266 to AAB16340 are

```

```

CC proteins encoded by some of the polynucleotide sequence given in the
CC present invention
XX
XX Sequence 316 AA;
Qy      Query Match      73.6%; Score 1185.5; DB 3; Length 316;
Db      Best local Similarity 71.9%; Pred. No. 1.9e-100;
Db      Matches 228; Conservative 38; Mismatches 50; Indels 1; Gaps 1;

Qy      1 MATLKVSDSVAPASDDAEQLRTAFEGWGTMEDLIISILHRSAPQRKVIQAHYETGED 60
Db      1 MSTLTVPOLPLPVADDCQDLRTAFAGWGTMEDLIISILHRSAPQRKVIQAHYETGED 60
Qy      61 LKLTLDKELSNDFERRAILLWTLPEGERDALLANFATKRTWSSNOVLMEVACTRTSTQLH 120
Db      61 LKLTLDKELSNDFERRAILLWTLPEGERDALLANFATKRTWSSNOVLMEVACTRTSTQLH 120
Qy      121 AROAYHARYKKSLEEDVAHHTTGDFRKLIVSLVTSYRYEGDEVNMTLAKQBAKLVHEKIK 180
Db      121 AROAYHARYKKSLEEDVAHHTTGDFRKLIVSLVTSYRYEGDEVNMTLAKQBAKLVHEKIK 180
Qy      181 DKHYNDESVIRILSTRSKAQINATFNRYODHGEELIKSLSEGGDDDKFLALRSTIOCL 240
Db      181 EKAYGHEDLIRILATRSKAQVNAATLNHYKBERGNDINKDLKT-DPKDAFLTLIRATVKCL 239
Qy      241 TRPELYFVDVLRSAINKGTDEGALTRIVTTTAEIDLVKVIAGEYQRNSIPLEKAITKDT 300
Db      240 TRPEKYEKVLRLAIKRGTDGALTRVAVTAEVDMKFISEYQRNSIPLDRAIVKDT 299
Qy      301 RGDYERKMLVALLGEDDA 317
Db      300 TGDYERKMLVALLIGHVEA 316

RESULT 8
AAB16321
ID AAB16321 standard; protein; 316 AA.
AC AAB16321;
XX
XX 31-OCT-2000 (first entry)
XX
XX pinus radiata amexin protein sequence SEQ ID NO:114.
XX
XX Eucalyptus grandis; pinus radiata; Monterey pine; modification;
XX plant cell wall; polysaccharide; polysaccharide biosynthetic pathway;
XX transgenic plant.
XX
XX pinus radiata.
XX
XX WO200022092-A2.
XX
XX 20-APR-2000.
XX
XX 08-OCT-1999; 99WO-NZ000169.
XX
XX 13-OCT-1998; 98US-00170862.
XX
XX 11-AUG-1999; 99US-0148426P.
XX
XX (GENE-) GENESIS RES & DEV CORP LTD.
XX (FLET-) FLETCHER CHALLENGE FORESTS LTD.
XX
XX Bloksberg LN;
XX
XX WPI; 2000-339328/29.
XX
XX N-PSDB; AAB67128.
XX
XX New genes encoding proteins involved in a plant polysaccharide
XX biosynthetic pathway, useful for modulating or altering the
XX polysaccharide content, composition or structure of the plant.
XX
XX Claim 17, Page 94-95; 301pp; English.
XX

```





PF 14-AUG-2002; 2002US-00219220.  
XX  
PR 04-JUN-1999; 99US-00325932.  
XX  
XX (GENE-) GENESIS RES & DEV CORP LTD.  
XX  
PI Plinn B, Lasham A;  
XX  
DR WPI; 2003-786916/74.  
XX  
PT New isolated polynucleotide useful for modulating programmed cell death,  
PT altering the development cycle of plant cells, and subsequently modifying  
PT plant development.  
XX  
PS Claim 12; Page 47-48; 21app; English.  
XX  
XX The invention describes an isolated polynucleotide (1) comprising a  
CC nucleotide sequence that is one of 145 fully defined sequences of 221-  
CC 3415 base pairs (bp), given in the specification, and/or its complements,  
CC reverse complements, reverse sequences, or sequences having 75, 90 or 95  
CC % sequence identity to or that hybridise under stringent hybridisation  
CC conditions to one of the 145 sequences. The methods and compositions of  
CC the present invention to do with (1) are useful for modulating programmed  
CC cell death and thereby altering the development cycle of plant cells, and  
CC altering plant development. This is the amino acid sequence of a protein  
CC associated with the programmed cell death pathway.  
XX  
SQ Sequence 316 AA;  
Query Match 72.7%; Score 1171.5; DB 7; Length 316;  
Best Local Similarity 71.3%; Pred. No. 3.8e-99;  
Matches 226; Conservative 39; Mismatches 51; Indels 1; Gaps 1;  
QY 1 MATLKVS DSPAPSDDAEQLRTAFEGWGTNEDLIISILAHSAEQRKVIRQAYHETYGED 60  
DB 1 MSTLTVPQPLPPVADDEQLRTAFAGWGTNEKLIISILHRRNAQRKIRQYAEFTYGED 60  
QY 61 LKTLIDKELSNDFERRALLMTLEPGRDALLANEAATKRTSSNOVMEVACTRTSTQOLH 120  
DB 61 LKRALDKELSNDFERRALLMTLEPGRDALLANEAATKRTSSNOVMEVACTRTSTQOLH 120  
QY 121 AROAYHARYKKSLEEDVAHHTTGDPRKLLVSLVTSYREGDEVNMTLAKQBAKLVHEKIK 180  
DB 121 AROAYHARYKKSMEDEVVAHHTTGDPRKLLVPLSSYRNDDEVNMTLAKQBAKLVHEKIS 180  
QY 181 DKHYNDEVDIRILSTRSKAQINATFNRYQDHEGEEILKSLSEGGDDDKFLALLRSTIOCL 240  
DB 181 EKAYGHEDLIRILSTRSKAQVNATFNRYKNEFGNDINKDLKT-DPKDAFLTIIRATVYKCL 239  
QY 241 TREPELVFVDVLRSAINTGTDEGALTRIVTTAEIDKLVIGEYQRNSIPLEKAITKOT 300  
DB 240 TREPEKTFEKVLRILAINRGITDEGALTRIVATRAEVDMKFISEEYQRNSIPLDRAIVKOT 299  
QY 301 RGDYERKMLVALLGEDDA 317  
DB 300 TGDYERKMLLALIGHVEA 316  
RESULT 11  
ADB94896  
ID ADB94896 standard; protein; 316 AA.  
XX  
XX ADB94896;  
XX  
XX 04-DEC-2003 (first entry)  
XX  
XX Programmed cell death pathway protein annexin #6.  
XX  
XX Programmed cell death; plant development; plant cell cycle; ATL2; DAD1;  
XX Dnaase; IIs; IIsd; nucellin-like aspartic protease; annexin; prohibitin;  
XX Ien-like protein; rac2; retinoblastoma-related protein; SIVA; TFIID;  
XX TEGT; xylogenic RNase; pur-alpha; cyteine protease; RPP5-like protein;  
XX GP 91 NADPH oxidase subunit; NPR-like protein; BAG-1;

KM defender against cell death; lethal leaf spot; lesion stimulating death;  
KM seven in absentia; transcription initiation factor;  
KM testis enhanced gene transcript.  
XX  
XX OS Eucalyptus grandis.  
XX  
XX PN US2003082724-A1.  
XX  
XX PD 01-MAY-2003.  
XX  
XX PF 14-AUG-2002; 2002US-00219220.  
XX  
XX PR 04-JUN-1999; 99US-00325932.  
XX  
XX XX (GENE-) GENESIS RES & DEV CORP LTD.  
XX  
XX PI Plinn B, Lasham A;  
XX  
XX DR WPI; 2003-786916/74.  
XX  
XX DR N-PSDB; ADB94854.  
XX  
XX PT New isolated polynucleotide useful for modulating programmed cell death,  
PT altering the development cycle of plant cells, and subsequently modifying  
PT plant development.  
XX  
PS Claim 12; Fig 27; 21app; English.  
XX  
XX The invention describes an isolated polynucleotide (1) comprising a  
CC nucleotide sequence that is one of 145 fully defined sequences of 221-  
CC 3415 base pairs (bp), given in the specification, and/or its complements,  
CC reverse complements, reverse sequences, or sequences having 75, 90 or 95  
CC % sequence identity to or that hybridise under stringent hybridisation  
CC conditions to one of the 145 sequences. The methods and compositions of  
CC the present invention to do with (1) are useful for modulating programmed  
CC cell death and thereby altering the development cycle of plant cells, and  
CC altering plant development. This is the amino acid sequence of a protein  
CC associated with the programmed cell death pathway.  
XX  
SQ Sequence 316 AA;  
Query Match 70.3%; Score 1132.5; DB 7; Length 316;  
Best Local Similarity 68.8%; Pred. No. 1.5e-95;  
Matches 218; Conservative 48; Mismatches 50; Indels 1; Gaps 1;  
QY 1 MATLKVS DSPAPSDDAEQLRTAFEGWGTNEDLIISILAHSAEQRKVIRQAYHETYGED 60  
DB 1 MATIAVPPSPVSPAPSDAEQLRTAFAGWGTNEDLIISILHRRNAQRKIRQYAEFTYGED 60  
QY 61 LKTLIDKELSNDFERRALLMTLEPGRDALLANEAATKRTSSNOVMEVACTRTSTQOLH 120  
DB 61 LKRALDKELSNDFERRALLMTLEPGRDALLANEAATKRTSSNOVMEVACTRTSSMELPM 120  
QY 121 AROAYHARYKKSLEEDVAHHTTGDPRKLLVSLVTSYREGDEVNMTLAKQBAKLVHEKIK 180  
DB 121 VROAYHARYKKSLEEDIAHHTTGDPRKLLVPLASTFRYGPVNVMTLARSSEKILHEKH 180  
QY 181 DKHYNDEVDIRILSTRSKAQINATFNRYQDHEGEEILKSLSEGGDDDKFLALLRSTIOCL 240  
DB 181 EKAYNHDELIRILSTRSKAQVNATFNRYKNEFGNAINKDL-KADPDEFLKILRSKICL 239  
QY 241 TREPELVFVDVLRSAINTGTDEGALTRIVTTAEIDKLVIGEYQRNSIPLEKAITKOT 300  
DB 240 TPEPEKTFEKVLRILAINRGITDEGALTRIVATRAEVDKRIKQYHRRNSVPLDRAIANDT 299  
QY 301 RGDYERKMLVALLGEDDA 317  
DB 300 SGDYERKMLLALVGHEDA 316  
RESULT 12  
AAG09201  
ID AAG09201 standard; protein; 211 AA.  
XX

AC AAG09201;  
XX  
DT 17-OCT-2000 (first entry)  
XX Arabidopsis thaliana protein fragment SEQ ID NO: 7040.  
XX  
KM Protein identification; signal transduction pathway; metabolic pathway;  
KM hybridisation assay; genetic mapping; gene expression control; promoter;  
KM termination sequence.  
XX Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-00301439.  
XX  
XX 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 23-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135533P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137224P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.

PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 26-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144633P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.

PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152353P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 26-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161922P.  
PR 28-OCT-1999; 99US-0161993P.  
PR 29-OCT-1999; 99US-0162142P.

Query Match 66.5%; Score 1071; DB 3; Length 211;  
Best Local Similarity 100.0%; Pred. No. 3.8e-90;  
Matches 211; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 107 MBVACTRTSTQLHARQAYHARYKKSLEEDVAHHTTDPFKLLVSLVTSYRYEGDEVNMT 166  
DB 1 MBVACTRTSTQLHARQAYHARYKKSLEEDVAHHTTDPFKLLVSLVTSYRYEGDEVNMT 60  
QY 167 LAQOAKLVHEKIKDKKYNDEVDVIRIISTSKAQINATFNRYODDHGEETILKSLERGGDD 226  
DB 61 LAQOAKLVHEKIKDKKYNDEVDVIRIISTSKAQINATFNRYODDHGEETILKSLERGGDD 120  
QY 227 DKFLALRSTIOCTTRELFFVDLRSAINKTGDBGALTRITYTTTAEIDLKVIAGEFYOR 286  
DB 121 DKFLALRSTIOCTTRELFFVDLRSAINKTGDBGALTRITYTTTAEIDLKVIAGEFYOR 180  
QY 287 RNSIPIEKAITTKOTRGDYERKMLVALLGEDDA 317  
DB 181 RNSIPIEKAITTKOTRGDYERKMLVALLGEDDA 211

RESULT 13  
ID AAG08875  
AAG08875 standard; protein; 211 AA.  
XX  
AC AAG08875;  
XX  
DT 17-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 6583.  
XX  
XX  
KM Protein identification; signal transduction pathway; metabolic pathway;  
KM hybridisation assay; genetic mapping; gene expression control; promoter;  
KM termination sequence.  
OS Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-00301439.  
XX  
PR 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 23-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0131499P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135533P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.



Qy 287 NSIPIEKATTKDTRGDEKMLVALLGDDA 317  
Db 181 NSIPIEKATTKDTRGDEKMLVALLGDDA 211

RESULT 14  
AAC06311  
ID AAC06311 standard; protein: 211 AA.

AC AAC06311;  
XX  
XX 17-OCT-2000 (first entry)  
XX  
XX Arabidopsis thaliana protein fragment SEQ ID NO: 3040.  
XX  
XX Protein identification; signal transduction pathway; metabolic pathway;  
KM hybridisation assay; genetic mapping; gene expression control; promoter;  
XX termination sequence.  
OS Arabidopsis thaliana.  
XX  
XX EP1033405-A2.  
XX  
XX 06-SEP-2000.  
XX  
XX 25-FEB-2000; 2000EP-0030139.  
XX  
XX 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 23-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.

PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-014187P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142290P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144684P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.

```
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149930P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151030P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161202P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.
```

Query Match 66.5%; Score 1071; DB 3; Length 211;

Best Local Similarity 100.0%; Pred.No.3.8e-90;

Matches 211; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 107 MEVACTRTSTQTLHAQAQYARYKSLSEEDVAHHTTGDPFRKLLVSLVTSYRYGDEVMNT 166
DB 1 MEVACTRTSTQTLHAQAQYARYKSLSEEDVAHHTTGDPFRKLLVSLVTSYRYGDEVMNT 60
QY 167 LAKQAKLVHEKIKDKHYNDEDEVIRILSTRSKAQINATFNRYODDHGEIILKSLSEGGDD 226
```

```
DB 61 LAKQAKLVHEKIKDKHYNDEDEVIRILSTRSKAQINATFNRYODDHGEIILKSLSEGGDD 120
QY 227 DKFLALRSTIOCLRPPELYFVDVRSAINKTGTDEGALTRIVTTRAFIDLVIGEYOR 286
DB 121 DKFLALRSTIOCLRPPELYFVDVRSAINKTGTDEGALTRIVTTRAFIDLVIGEYOR 180
QY 287 RNSIPLEKAITKDTRGDYKMKLVALLGEDDA 317
DB 181 RNSIPLEKAITKDTRGDYKMKLVALLGEDDA 211

RESULT 15
AAG23493
ID AAG23493 standard; protein; 317 AA.
XX
AC AAG23493;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 26821.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0133256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
```

```
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 22-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142059P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144355P.
PR 19-JUL-1999; 99US-0144351P.
PR 19-JUL-1999; 99US-0144352P.
PR 19-JUL-1999; 99US-0144353P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144652P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 03-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.

PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0152363P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 66.2%; Score 1067; DB 3; Length 317;
Best local Similarity 64.0%; Pred. No. 1,6e-89;
Matches 203; Conservative 50; Mismatches 64; Indels 0; Gaps 0;
```



```
Db      1 MASLKVPSNVPLPEDDAEQHKAFGSGWGTNEKLIISILAHRNAAQRSLIRSVYAATYNED 60
QY      61 LKTLDKETLNDPERAILMTLEPGERDALLANEATKWTSSNOVIMEVACTRTSTQLH 120
Db      61 LKALDKELSSDERAVMLTLDPPERDAYLAKESTKMTKNNWVLVEIACTRPALELIK 120
QY      121 ARQAYHARYKKSLEEDVAHHTTGDPRKLVSLVTSYRYEGDEVNMTLAKOEAKLVHEKIK 180
Db      121 VKQAYQARYKKSIEEDVAQHTSGDLKLLPLVSTFRYEGDDVNMMLARSBAKILHEKVS 180
QY      181 DKRYNDEVIIRIISTRSKAQINATFNRYODDHGEIILKSLEEGDDDKFLALLRSTIQCL 240
Db      181 EKSYSDDDFIRILITRSKAQIGATLNHNNEYGNAINKNLKESDDNDYMKLLRAVITCL 240
QY      241 TRPELYFVDVLRSAINKTGTDEGALTRITTRAIDLKVIGEYQRRNSIPLEKATKDT 300
Db      241 TYPEKHPEKVLRLISINKMGTDWGLTRVVTTRTEVDMERIKEEYQRRNSIPLDRAIAKDT 300
QY      301 RGDYERKLVALLGEDDA 317
Db      301 SGDYEDMLVALLGHGDA 317
```

Search completed: August 19, 2005, 20:40:41  
Job time : 169 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 19, 2005, 20:36:40 ; Search time 40 Seconds  
(without alignments)  
762.518 Million cell updates/sec

Title: US-10-690-564-2

Perfect score: 1611

Sequence: 1 MATLKVSDSVAPASDDAEQL.....KDTRGDYKKLVALLGEDDA 317

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0 %

Maximum Match 100 %

Listing first 45 summaries

Database :

1: PIR.79:\*  
2: PIR1:\*  
3: PIR2:\*  
4: PIR3:\*  
5: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1611	100.0	317	2	C86479
2	1202.5	74.6	316	2	T31428
3	1128.5	70.0	315	2	T10805
4	1117.5	69.4	314	2	S66274
5	1107	68.7	315	2	T10807
6	1080	67.0	315	2	T06322
7	1067.5	66.3	308	2	T09552
8	1037.5	64.4	316	2	T50027
9	1000.5	62.1	314	2	T02961
10	991.5	61.5	318	2	T50026
11	949.5	58.9	314	2	T02975
12	772.5	48.0	271	2	S56674
13	772.5	37.4	321	2	A84809
14	523	32.5	316	2	B96704
15	501	31.1	316	1	L0H015
16	496	30.8	316	2	A57076
17	496	30.8	357	2	B57076
18	472	29.3	323	1	L0H015
19	462	28.7	319	2	H84808
20	461.5	28.6	673	1	A0H068
21	454	28.2	324	1	LUR73
22	451	28.0	505	2	A5152
23	450	27.9	503	1	L0B011
24	450	27.9	505	1	S23447
25	448.5	27.8	673	1	S01786
26	447	27.7	488	1	L0H017
27	444	27.6	503	1	LURB11
28	435.5	27.0	318	1	LURP4
29	434	26.9	463	2	S29170

30	429.5	26.7	673	1	S52844	annexin VI - rat
31	425.5	26.4	321	1	A42077	annexin IV - human
32	424.5	26.4	671	2	UC2029	annexin - chicken
33	421.5	26.2	319	1	L0B04	annexin IV - bovin
34	413.5	25.7	321	1	LURF10	annexin X - fruit
35	408.5	25.4	512	2	S70644	annexin VII - Afri
36	407.5	25.3	321	1	LURC5	annexin V - chicke
37	405	25.1	340	1	JQ1297	annexin II type 1
38	403.5	25.0	339	1	L0B036	annexin II - bovin
39	402	25.0	339	1	LURH2	annexin II - chick
40	398.5	24.7	339	1	LURH36	annexin II - human
41	398.5	24.7	340	1	JQ1298	annexin II type 2
42	396.5	24.6	339	1	LURM36	annexin II - mouse
43	396	24.6	346	1	LURF1	annexin I - rat
44	393	24.4	346	1	LURH	annexin I - human
45	392.5	24.4	327	1	LURH8	annexin VIII - hum

#### ALIGNMENTS

##### RESULT 1

C86479

probable annexin protein - Arabidopsis thaliana

C/Species: Arabidopsis thaliana (mouse-ear crests)

C/Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004

C/Accession: C86479

R/Theologian, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.

ansen, N.F.; Hughes, B.; Hultzer, L.

Nature 408, 816-820, 2000

A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luo, J.S.; Malt, R.; Marzalli

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shum, P.; Southwick, A.M.; Sun, H.; Tallon,

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A/Reference number: A86141; MUID:21016719; PMID:11130712

A/Accession: C86479

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-317 <STO>

A/Cross-references: UNIPROT:Q9SYT0; GB:AE005172; NID:98778967; PIDN:AA79882.1; GSPDB:G

C/Genetics:

A/Map position: 1

C/Superfamily: annexin I; annexin repeat homology

Query Match 100.0%; Score 1611; DB 2; Length 317;  
Best Local Similarity 100.0%; Pred. No. 1.6e-97;  
Matches 317; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MATLKVSDSVAPASDDAEQLRTAFEGWGTNEDLISILHARSAPORRVIAQAHETGYED	60
DB	1	MATLKVSDSVAPASDDAEQLRTAFEGWGTNEDLISILHARSAPORRVIAQAHETGYED	60
QY	61	LKLTDLKELSNDFERRILLMTLEPGERDALLANEATKRWTSNQVIMEVACTRTSTQLH	120
DB	61	LKLTDLKELSNDFERRILLMTLEPGERDALLANEATKRWTSNQVIMEVACTRTSTQLH	120
QY	121	ARQAHARVKKSLSEBVVAHHTTDPFKLVSLVTSYRSGEDVNMTLAKQEAALVHEKIK	180
DB	121	ARQAHARVKKSLSEBVVAHHTTDPFKLVSLVTSYRSGEDVNMTLAKQEAALVHEKIK	180
QY	181	DKHYNDEVDVIRILSTRSKAQINATFNRYODDHGEELIKSLSEGGDDDKFLALRSTIQL	240
DB	181	DKHYNDEVDVIRILSTRSKAQINATFNRYODDHGEELIKSLSEGGDDDKFLALRSTIQL	240
QY	241	TRELYFVDVLRSAINKTGTDEGALTRIVTTAEIDLKVIAGEBYQRNSIPEKAITKOT	300
DB	241	TRELYFVDVLRSAINKTGTDEGALTRIVTTAEIDLKVIAGEBYQRNSIPEKAITKOT	300
QY	301	RGDYKKLVALLGEDDA 317	
DB	301	RGDYKKLVALLGEDDA 317	

Db 301 RGDYKMLVALLGEDDA 317

## RESULT 2

TJ1428

fiber annexin - upland cotton

C/Species: Gossypium hirsutum (upland cotton)

C/Date: 29-Oct-1999 #sequence\_rev15000 #text\_change 09-Jul-2004

C/Accession: TJ1428

R/Shin, H.; Brown, R.M.

submitted to the EMBL Data Library, February 1997

A/Reference number: Z21028

A/Accession: TJ1428

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: mRNA

A/Residues: 1-316 <SH1>

A/Cross-references: UNIPROT:P93157; EMBL:U089609; NID:g3493171; PID:g3493172

A/Experimental source: strain Texas Marker1

C/Superfamily: annexin I; annexin repeat homology

C/Keywords: calcium binding

Query Match 74.6%; Score 1202.5; DB 2; Length 316;

Best Local Similarity 73.5%; Pred. No. 5.3e-71;

Matches 230; Conservative 37; Mismatches 45; Indels 1; Gaps 1;

Db 1 MATLKVSQSVAPSPDADAEQRTAFEGMGNTEDLIISLAHSAEQRKVIROAYHETGSD 60

1 MATLVPTVPSVSEDCQQRKAFSGMGNTGIIIDIGHRNAEQRLIRKTYAETGSD 60

QY 61 LKTLDELNDPERAILLMTLEPERDALLANEATKRWTSNOVLMVACTRTSTQLH 120

61 LKALDELNDPERAILLMTLEPERDALLANEATKRWTSNOVLMVACTRTSTQLH 120

QY 121 ARQAYHARYKKSLEEDVAHHTTGDPRKLLVSLVTSRYEGDEVNMTLAKOEAALVHEKIK 180

121 ARQAYHARYKKSLEEDVAHHTTGDPRKLLVSLVTSRYEGDEVNMTLAKOEAALVHEKIK 180

QY 121 ARQAYHARYKKSLEEDVAHHTTGDPRKLLVSLVTSRYEGDEVNMTLAKOEAALVHEKIK 180

121 ARQAYHARYKKSLEEDVAHHTTGDPRKLLVSLVTSRYEGDEVNMTLAKOEAALVHEKIK 180

QY 181 DKRYNDEVIRILSTRSKAQINATFNRYODDHEGELIKSLBEGDDDDKFLALRSTIOCT 240

181 DKRYNDEVIRILSTRSKAQINATFNRYODDHEGELIKSLBEGDDDDKFLALRSTIOCT 240

QY 241 TRPELVFVDVLRSAINKGTDEGALTRIVTTTRAEIDLKVIIGEYQRNSIPEKAITKDT 300

240 VPEKFEKVLRLAIRRGTDGALTRIVCTRAEVDLKIADBYQRNSVPLTRAIKDT 299

QY 301 RGDYKMLVALLG 313

300 RGDYKMLVALLG 312

## RESULT 3

T10805

annexin - upland cotton (fragment)

C/Species: Gossypium hirsutum (upland cotton)

C/Date: 16-Jul-1999 #sequence\_rev15000 #text\_change 09-Jul-2004

C/Accession: T10805

R/Potikha, T.S.; Delmer, D.P.

submitted to the EMBL Data Library, October 1996

A/Reference number: Z17161

A/Accession: T10805

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: mRNA

A/Residues: 1-315 <POT>

A/Cross-references: UNIPROT:P93157; EMBL:U073746; NID:g1843524; PID:g1843525

A/Experimental source: strain Acala SJ-2; cotton fiber

C/Genetics:

A/Note: Anngh1

C/Superfamily: annexin I; annexin repeat homology

C/Keywords: calcium binding

F:85-156/Domain: annexin repeat homology <AXR>

Query Match 70.0%; Score 1128.5; DB 2; Length 315;

Best Local Similarity 69.2%; Pred. No. 3.4e-66;

Matches 216; Conservative 40; Mismatches 55; Indels 1; Gaps 1;

QY 2 ATLKVSDSVAPSPDADAEQRTAFEGMGNTEDLIISLAHSAEQRKVIROAYHETGSD 61

1 ATLKVPTVPSVSEDCQQRKAFSGMGNTGIIIDIGHRNAEQRLIRKTYAETGSD 60

QY 62 LKTLDELNDPERAILLMTLEPERDALLANEATKRWTSNOVLMVACTRTSTQLH 121

61 LKALDELNDPERAILLMTLEPERDALLANEATKRWTSNOVLMVACTRTSTQLH 120

QY 122 RQAYHARYKKSLEEDVAHHTTGDPRKLLVSLVTSRYEGDEVNMTLAKOEAALVHEKIK 181

121 RQAYHARYKKSLEEDVAHHTTGDPRKLLVSLVTSRYEGDEVNMTLAKOEAALVHEKIK 180

QY 182 KRYNDEVIRILSTRSKAQINATFNRYODDHEGELIKSLBEGDDDDKFLALRSTIOCT 241

181 KRYNDEVIRILSTRSKAQINATFNRYODDHEGELIKSLBEGDDDDKFLALRSTIOCT 239

QY 242 RPELVFVDVLRSAINKGTDEGALTRIVTTTRAEIDLKVIIGEYQRNSIPEKAITKDT 301

240 VPEKFEKVLRLAIRRGTDGALTRIVCTRAEVDLKIADBYQRNSVPLTRAIKDT 299

QY 302 RGDYKMLVALLG 313

300 RGDYKMLVALLG 311

## RESULT 4

S66274

annexin - pepper

C/Species: Capsicum annuum (pepper)

C/Date: 15-Feb-1997 #sequence\_rev15000 #text\_change 09-Jul-2004

C/Accession: S66274

R/Proust, J.; Houline, G.; Schantz, M.L.; Schantz, R.

FEBS Lett. 383, 208-212, 1996

A/Title: Characterization and gene expression of an annexin during fruit development in

A/Reference number: S66274; MUID:96198603; PMID:8925897

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-314 <PRO>

A/Cross-references: UNIPROT:Q42657; EMBL:X93308; NID:g1071659; PID:CA63710.1; PID:g107

A/Note: the sequence of residues 155-230 is shown in line for annexin (alfalfa) rather t

C/Superfamily: annexin I; annexin repeat homology

F:14-85/Domain: annexin repeat homology <AX1>

F:86-157/Domain: annexin repeat homology <AX2>

F:169-236/Domain: annexin repeat homology <AX3>

F:242-313/Domain: annexin repeat homology <AX4>

Query Match 69.4%; Score 1117.5; DB 2; Length 314;

Best Local Similarity 67.1%; Pred. No. 1.7e-65;

Matches 212; Conservative 58; Mismatches 43; Indels 3; Gaps 1;

QY 1 MATLKVSQSVAPSPDADAEQRTAFEGMGNTEDLIISLAHSAEQRKVIROAYHETGSD 60

1 MATLVPTVPSVSEDCQQRKAFSGMGNTGIIIDIGHRNAEQRLIRKTYAETGSD 60

QY 61 LKTLDELNDPERAILLMTLEPERDALLANEATKRWTSNOVLMVACTRTSTQLH 120

61 LKALDELNDPERAILLMTLEPERDALLANEATKRWTSNOVLMVACTRTSTQLH 120

QY 121 ARQAYHARYKKSLEEDVAHHTTGDPRKLLVSLVTSRYEGDEVNMTLAKOEAALVHEKIK 180

121 ARQAYHARYKKSLEEDVAHHTTGDPRKLLVSLVTSRYEGDEVNMTLAKOEAALVHEKIK 180

QY 181 DKRYNDEVIRILSTRSKAQINATFNRYODDHEGELIKSLBEGDDDDKFLALRSTIOCT 240

181 DKRYNDEVIRILSTRSKAQINATFNRYODDHEGELIKSLBEGDDDDKFLALRSTIOCT 237

QY 241 TRPELVFVDVLRSAINKGTDEGALTRIVTTTRAEIDLKVIIGEYQRNSIPEKAITKDT 300

240 VPEKFEKVLRLAIRRGTDGALTRIVCTRAEVDLKIADBYQRNSVPLTRAIKDT 297



Db 63 SSDPEKAVLLMTLDPAREDAFLANQATKMLTNNSTIVEIASRSPLEILKAKQAVQVR 122  
QY 130 KSLSEEDVAHHTGDPKRLVSVTSRYRGDEVNMTLAKQEKLVHEKIDGNYDEVD 189  
Db 123 KSLSEEDVAHHTGDKRLKPLVGVGHRVYRGDEVNMTLASEAKLHEKIDADVAHMDL 182  
QY 190 IRLSTRSKAQINATFNRYODDGEELIKSLSEGGDDDKFLALRSTIOCTRLPELVFD 249  
Db 183 IRLVTRSKAQINATLNHYNNRGVNDKDLFT-DSDDEVLKILRAIKGLTYPEKTFEE 241  
QY 250 VLRSAINKTGTDEGALTRIVTTTAEIDLKVIIGEYGRNSIPLEKAITKOTRDYERMTV 309  
Db 242 LRLALINKMGTDENALTRVVTTRAEDVLQRIAEYGRNSVPLDRAIDKOTSGDYQKILL 301  
QY 310 ALLGDEDD 316  
Db 302 ALMGHDE 308

## RESULT 8

T50027  
annexin-like protein - Arabidopsis thaliana  
N/Alternate names: protein T31P16.220  
C/Species: Arabidopsis thaliana (mouse-ear cress)  
C/Date: 02-Jun-2000 #sequence\_revision 02-Jun-2000 #text\_change 09-Jul-2004  
C/Accession: T50027  
R/Bevan, M.; Zimmermann, W.; Grueniseen, A.; Wambutt, R.; Kalicki, J.; Wohlmann, P.; St  
Submitted to the Protein Sequence Database, May 2000  
A/Reference number: Z25027  
A/Accession: T50027  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-316 <BEV>  
A/Cross-references: UNIPROT:Q9LX07; EMBL:AL356332; GSPDB:GNO0063; ATSP:T31P16.220  
A/Experimental source: cultivar Columbia; BAC clone T31P16  
C/Genetics:  
A/Gene: ATSP:T31P16.220  
A/Map position: 5  
A/Intons: 26/1; 74/3; 147/3  
C/Superfamily: annexin I; annexin repeat homology

Query Match 64.4%; Score 1037.5; DB 2; Length 316;  
Best Local Similarity 63.1%; Pred. No. 2.7e-60;  
Matches 200; Conservative 55; Mismatches 61; Indels 1; Gaps 1;  
QY 1 MATLKVSDSVPAESDDAEQLRTAFEGWGTNEDLIISILHRSABQRKVIROAVHETYGED 60  
Db 1 MASLKVPATVPPEEDAEOLYKAFKGTNERMISILHARNATORSFIRAVVAYANYNKD 60  
QY 61 LRLTLDEKLSNDFERRAILMTLEPGERDALLANEATKRTSSNOVLMVEVACTRTSTQLH 120  
Db 61 LRLTLDEKLSNDFERRAILMTLEPGERDALLANEATKRTSSNOVLMVEVACTRTSTQLH 120  
QY 121 ARQAVHARYKKSLEEDVAHHTGDPKRLVSVTSRYRGDEVNMTLAKQEKLVHEKID 180  
Db 121 AKQAVHARYKKSLEEDVAHHTGDPKRLVSVTSRYRGDEVNMTLAKQEKLVHEKID 180  
QY 181 DKRYNDEVIRILSTRSKAQINATFNRYODDGEELIKSLSEGGDDDKFLALRSTIOCT 240  
Db 181 EKAIVADDDILIRILSTRSKAQISATLNHYNNRGVNDKDLFT-DSDDEVLKILRAIKGLTYPEKTFEE 241  
QY 241 TRPELVFVNLRSAINKTGTDEGALTRIVTTTAEIDLKVIIGEYGRNSIPLEKAITKOT 300  
Db 240 TYPEKTFEKLQAINLGLTDENGILTRVITTRAEPFMERIKERYIRNSVPLDRAIAKOT 299  
QY 301 RGDYERKMLVALLGEDDA 317  
Db 300 HGDYEDILALLIGHDA 316

## RESULT 9

annexin P33 - maize  
C/Species: Zea mays (maize)  
C/Date: 24-Mar-1999 #sequence\_revision 24-Mar-1999 #text\_change 09-Jul-2004  
C/Accession: T02961  
R/Pattey, N.H.; James, N.C.; Greenland, A.J.  
Plant Physiol. 112, 1391-1396, 1996  
A/Title: CDNA isolation and gene expression of maize annexins P33 and P35.  
A/Reference number: Z14796; MUID:97092863; PMID:8938425  
A/Accession: T02961  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: mRNA  
A/Residues: 1-314 <BAT>  
A/Cross-references: UNIPROT:Q43863; EMBL:X98244; PIDN:CAA66900.1  
A/Experimental source: cultivar clipper; root tip  
C/Superfamily: annexin I; annexin repeat homology  
P/14-85/Domain: annexin repeat homology <XLR>  
Query Match 62.1%; Score 1000.5; DB 2; Length 314;  
Best Local Similarity 60.3%; Pred. No. 6.8e-58;  
Matches 190; Conservative 53; Mismatches 71; Indels 1; Gaps 1;  
QY 1 MATLKVSDSVPAESDDAEQLRTAFEGWGTNEDLIISILHRSABQRKVIROAVHETYGED 60  
Db 1 MASLKVPATVPPEADCDQLRAFOGWGTNEALISILHRDAQRALRAVAAEYGE 60  
QY 61 LRLTLDEKLSNDFERRAILMTLEPGERDALLANEATKRTSSNOVLMVEVACTRTSTQLH 120  
Db 61 LRLSTIDELSGFERAVIIMTLDPAREDAVLNEARKKPKPNRVLVEICTRTSAQIRA 120  
QY 121 ARQAVHARYKKSLEEDVAHHTGDPKRLVSVTSRYRGDEVNMTLAKQEKLVHEKID 180  
Db 121 TRQAVHERPKRSLEEDIAHVTGDPFKLVLPLVSTYRGDPVNTRLASEAKLHEKID 180  
QY 181 DKRYNDEVIRILSTRSKAQINATFNRYODDGEELIKSLSEGGDDDKFLALRSTIOCT 240  
Db 181 HRAVSDDELIRILTRSKAQILATFNRYODDGEELIKSLSEGGDDDKFLALRSTIOCT 239  
QY 241 TRPELVFVNLRSAINKTGTDEGALTRIVTTTAEIDLKVIIGEYGRNSIPLEKAITKOT 300  
Db 240 SCPDRYFEKVAQALAGLTDENSLTRVITTRAEDVLKIKAYGRNSVRLERAVAGDT 299  
QY 301 RGDYERKMLVALLGED 315  
Db 300 SGDYESMLALLGOE 314  
RESULT 10  
T50026  
annexin-like protein - Arabidopsis thaliana  
N/Alternate names: protein T31P16.210  
C/Species: Arabidopsis thaliana (mouse-ear cress)  
C/Date: 02-Jun-2000 #sequence\_revision 02-Jun-2000 #text\_change 09-Jul-2004  
C/Accession: T50026  
R/Bevan, M.; Zimmermann, W.; Grueniseen, A.; Wambutt, R.; Kalicki, J.; Wohlmann, P.; St  
Submitted to the Protein Sequence Database, May 2000  
A/Reference number: Z25027  
A/Accession: T50026  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-318 <BEV>  
A/Cross-references: UNIPROT:Q9LX08; EMBL:AL356332; GSPDB:GNO0063; ATSP:T31P16.210  
A/Experimental source: cultivar Columbia; BAC clone T31P16  
C/Genetics:  
A/Gene: ATSP:T31P16.210  
A/Map position: 5  
A/Intons: 26/1; 74/3; 147/3  
C/Superfamily: annexin I; annexin repeat homology  
Query Match 61.5%; Score 991.5; DB 2; Length 318;  
Best Local Similarity 60.8%; Pred. No. 2.6e-57;  
Matches 194; Conservative 52; Mismatches 70; Indels 3; Gaps 2;  
QY 1 MATLKVSDSVPAESDDAEQLRTAFEGWGTNEDLIISILHRSABQRKVIROAVHETYGED 60



```
QY 61 LKTLDKELSNDEPERAILLWTLTEPGERDALLAN-----EATKRWTSNNOVLMVEACCTRTS 115
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 LILVLSSELSDGMKAVSVSTTYTPARDALVNKILNKKKSLLENLKIYVISTCTTSP 120
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 116 TOLLHARQAVHARYKKSLEEDVAHHTTGDERKLLVSLVTSYRYEGDEVNMTLAKOBAKLV 175
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 121 NHLIAVRKAYCSLFDSLSBEHIASLPFLAKLLVTLASFYDOKORTDAEVATIBAMLI 180
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 176 HEKIKOKHNVDEVYIRILSTRSKAQINATNRQODHGEELIKSLSEGGDDDKFALLIIS 235
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 181 REAIEKKQDLHDHVLVILGTRSYOURETFVAKKXVGVITIDVDGCPEDADLRSLIKV 240
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 236 TICCLTRPELFPVDVRSAINKTGTDEGALTRIVTTRAEIDLKVIAGEYQRRNSIPLEKA 295
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 241 AIFCIDTPEHGFPAKVVRDSEEGRTBEDSLTRALVTRAEIDLKVKAGEYTNMTSTMDNA 300
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 296 ITKDTRGDYKMLVALIG 313
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 301 ITGDISGDYKDFITLLIG 318
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

## RESULT 14

```
B96704
Probable annexin T23K23.6 [imported] - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C/Accession: B96704
R/Theologian: A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chn, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Hlitzar, L.
Nature 408, 816-820, 2000
A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaytin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luo, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A/Reference number: A86141; MUID:21016719; PMID:11130712
A/Accession: B96704
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-316 <STO>
A/Cross-references: UNIPROT:Q9C9X3; GB:AE005173; NID:g6553887; PIDN:AAFL6553.1; GSPDB:GN
C/Genetics:
A/Gene: T23K23.6
C/Map position: 1
C/Superfamily: annexin I; annexin repeat homology
```

```
Query Match 32.5%; Score 523; DB 2; Length 316;
Best Local Similarity 34.7%; Pred. No. 6,7e-27;
Matches 109; Conservative 75; Mismatches 128; Indels 2; Gaps 2;

QY 1 MATLKVSQSVAPSDAEDQLRTAFEGWGTNEDLIISILARSAEQRKVIHQAYHETYGED 60
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 MATMKIPMTVPSPRVADQLFKAFKGTGCDTSYITINILAHNNATQRLALIQEYETKPSDD 60
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 61 LILVLSSELSDGMKAVSVSTTYTPARDALVNKILNKKKSLLENLKIYVISTCTTSP 120
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 LKTLDKELSNDEPERAILLWTLTEPGERDALLAN-----EATKRWTSNNOVLMVEACCTRTS 175
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 61 LKTLDKELSNDEPERAILLWTLTEPGERDALLAN-----EATKRWTSNNOVLMVEACCTRTS 120
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 LKTLDKELSNDEPERAILLWTLTEPGERDALLAN-----EATKRWTSNNOVLMVEACCTRTS 120
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 121 ARQAVHARYKKSLEEDVAHHTTGDERKLLVSLVTSYRYEGDEVNMTLAKOBAKLV 180
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 121 IKQVYSTPFGVLEEDIESASGNKKVLAAYLNTTRVGEPEIDNASVENDARTLSAVA 180
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 181 DKHYNED-VIRILSTRSKAQINATNRQODHGEELIKSLSEGGDDDKFALLIISRTIQC 239
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 181 RKKSDDQGLIQFTDRSRTHIVAVRSTYSWKGELKXIRP-ETRGNEHNLITLIQC 239
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 240 LTRPELFPVDVRSAINKTGTDEGALTRIVTTRAEIDLKVIAGEYQRRNSIPLEKAITKD 299
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 240 AENSCEFPAKALRKSMKGLGTDLTALIRIVTTRAEVDMQFIITEYRKRYKTYLNAVHSD 299
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 300 TRGDYKMLVALIG 313
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

```
Db 300 TTSHYRTFLSLILG 313
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

## RESULT 15

```
LUNUIS
annexin XII, intestinal [validated] - human
C/Species: Homo sapiens (man)
C/Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 03-Jul-2004
C/Accession: A41733; S18162
R/Wice, B.M.; Gordon, J.I.
J. Cell Biol. 116, 405-422, 1992
A/Title: A strategy for isolation of cDNAs encoding proteins affecting human intestinal
A/Reference number: A41733; MUID:92112982; PMID:1530946
A/Accession: A41733
A/Molecule type: mRNA
A/Residues: 1-316 <WIC>
A/Cross-references: UNIPROT:P27216; EMBL:Z11502; NID:g33979; PIDN:CAA77578.1; PID:g33980
C/Comment: Annexins undergo reversible, calcium-dependent binding to membrane phospholip
derstood.
C/Genetics:
A/Gene: GDB:ANXA13
A/Cross-references: GDB:9836021; OMIM:602573
A/Map position: 8q24.1-8q24.2
C/Superfamily: annexin I; annexin repeat homology
C/Keywords: blocked amino end; calcium binding; duplication; endonexin fold; lipoprotein
F/2-316/Product: annexin intestinal homology <AX1>
F/17-88/Domain: annexin repeat homology <AX1>
F/28-44/Region: endonexin fold #status predicted
F/89-160/Domain: annexin repeat homology <AX2>
F/100-116/Region: endonexin fold #status predicted
F/172-244/Domain: annexin repeat homology <AX3>
F/184-200/Region: endonexin fold #status predicted
F/248-316/Domain: annexin repeat homology <AX4>
F/259-275/Region: endonexin fold #status predicted
F/2/Modified site: myristylated amino end (Gly) (in mature form) #status experimental
F/2/Modified site: aspartic acid (Asn) #status predicted
```

```
Query Match 31.1%; Score 501; DB 1; Length 316;
Best Local Similarity 38.9%; Pred. No. 1.8e-25;
Matches 117; Conservative 61; Mismatches 115; Indels 8; Gaps 3;
```

```
QY 16 DAEQLRTAFEGWGTNEDLIISILARSAEQRKVIHQAYHETYGEDLTKTLDKELSNDEPER 75
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 19 DAKLKNKCKMGWGTNEAIIETLSGRTSPEROIKQKYATYKXLEBVLKSELGSPFK 78
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 76 AILWTLTEPGERDALLAN-----EATKRWTSNNOVLMVEACCTRTSQTLLHARQAVHARYKKSLE 135
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 79 TALALLDRPSEYAAARQLQKMKGLGTDSEVLIEFLCTRTNKEIIAIKEAYQRLPDRSLES 138
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 136 DVANHTTGPFRKLLVSLVTSYRYEGDEVNMTLAKOBAKLVHEKIDKHYNEDVI-RILS 194
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 139 DVKGDTSNKKILVSLQANRNEGDVDKDLADGADKLYAGGRTGTDLAFNEVLA 198
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 195 TTSKQAINATNRQODHGEELIKSLSE---GDDDDKFLALRSTIQCCTTRPELFPVDV 251
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 199 KASYKRLATFGAYQLLIGKIDEBARIEBTSDDLQKAYITLYRCQDQ---EDFAEEL 254
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 252 RSNINKTGTDEGALTRIVTTRAEIDLKVIAGEYQRRNSIPLEKAITKDTRGDYKMLVAL 311
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 255 YKSKAGAGDEFTLRIVVTRAEVLDQIKAKFQEKYQKSLSDWVRSPTSGDFRKLVAL 314
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 312 L 312
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 315 L 315
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

Search completed: August 19, 2005, 20:47:17  
Job time : 41 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 19, 2005, 20:09:22 ; Search time 43 Seconds

550.320 Million cell updates/sec

Title: US-10-690-564-2

Perfect score: 1611  
Sequence: 1 MATLKVSDSVPA~~PSD~~AEQL.....KOTRGDYEKMLVALGEDDA 317

Scoring table: BLOSUM62

Gapor 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%

Listing first 45 summaries

```
Issue_Patents_AA:*
1: /cgn2_6/prodata/1/iaa/5A.COMB:.rep:*
2: /cgn2_6/prodata/1/iaa/5B.COMB:.rep:*
3: /cgn2_6/prodata/1/iaa/6A.COMB:.rep:*
4: /cgn2_6/prodata/1/iaa/6B.COMB:.rep:*
5: /cgn2_6/prodata/1/iaa/PCtUS.COMB:.rep:*
6: /cgn2_6/prodata/1/iaa/backfill1.rep:*
```

**Pred. No.** is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1171.5	72.7	316	4	US-09-335-932A-62	Sequence 62, Applicant 1
2	866.5	53.8	239	4	US-09-325-932A-65	Sequence 65, Applicant 1
3	573	35.6	184	4	US-09-325-932A-66	Sequence 66, Applicant 1
4	472	29.3	323	4	US-08-948-276-5	Sequence 5, Applicant 1
5	472	29.3	323	4	US-09-949-016-6593	Sequence 6593, Applicant 1
6	472	29.3	334	4	US-09-949-016-10475	Sequence 10475, Applicant 1
7	461.5	28.6	681	4	US-09-949-016-10475	Sequence 10475, Applicant 1
8	461.5	28.6	681	4	US-09-949-016-10476	Sequence 10476, Applicant 1
9	451	28.0	557	4	US-09-949-016-7621	Sequence 7621, Applicant 1
10	450	27.9	503	3	US-08-526-136-2	Sequence 2, Applicant 1
11	450	27.9	505	3	US-08-526-136-4	Sequence 4, Applicant 1
12	449	27.9	111	4	US-09-325-932A-63	Sequence 63, Applicant 1
13	447	27.7	466	3	US-08-526-136-13	Sequence 13, Applicant 1
14	447	27.7	467	4	US-09-949-016-7070	Sequence 7070, Applicant 1
15	447	27.7	467	4	US-09-949-016-7071	Sequence 7071, Applicant 1
16	447	27.7	489	4	US-09-949-016-7068	Sequence 7068, Applicant 1
17	447	27.7	489	4	US-09-949-016-7069	Sequence 7069, Applicant 1
18	425.5	26.4	321	4	US-08-948-276-4	Sequence 4, Applicant 1
19	425.5	26.4	324	4	US-09-949-016-10531	Sequence 10531, Applicant 1
20	421.5	26.2	319	3	US-08-526-136-14	Sequence 14, Applicant 1
21	398.5	24.7	377	4	US-09-949-016-8302	Sequence 8302, Applicant 1
22	393	24.4	346	4	US-08-948-276-6	Sequence 6, Applicant 1
23	393	24.4	370	4	Sequence 11534, Applicant 1	Sequence 11534, Applicant 1
24	390	24.2	319	1	US-08-949-016-11534	Sequence 1, Applicant 1
25	390	24.2	320	1	US-08-135-746-1	Sequence 3, Applicant 1
26	390	24.2	320	4	US-08-135-746-3	Sequence 1, Applicant 1
27	390	24.2	320	6	US-08-948-276-1	Sequence 1, Applicant 1
					Patent No. 5225537-4	Patent No. 5225537-4

## ALIGNMENTS

28	390	24.2	320	6	5255537-4	Patent No. 5255537
29	390	24.2	323	4	US-09-949-016-10177	Sequence 10177, A
30	389	24.1	327	3	US-09-324-096A-2	Sequence 2, Appl.
31	389	24.1	327	3	US-09-324-096A-2	Sequence 4, Appl.
32	389	24.1	327	3	US-09-324-096A-6	Sequence 6, Appl.
33	389	24.1	327	3	US-09-970-969-4	Sequence 2, Appl.
34	389	24.1	327	4	US-09-970-969-4	Sequence 4, Appl.
35	389	24.1	327	4	US-09-970-969-6	Sequence 6, Appl.
36	389	24.1	327	4	US-09-970-969-6	Sequence 8, Appl.
37	376.5	23.4	319	2	US-08-949-276-3	Sequence 2, Appl.
38	376.5	23.4	319	2	US-08-949-276-3	Sequence 4, Appl.
39	374.5	23.2	319	4	US-08-949-276-2	Sequence 2, Appl.
40	374.5	23.2	357	4	US-09-949-016-8624	Sequence 8624, Appl.
41	370	21.9	421	4	US-09-0270-767-11904	Sequence 11904, A
42	352.5	21.9	324	4	US-09-010-147B-16	Sequence 16, Appl.
43	330	20.5	73	4	US-09-325-932A-64	Sequence 64, Appl.
44	330	20.5	73	4	US-09-325-932A-64	Sequence 64, Appl.
45	147	9.1	151	4	US-09-621-976-980	Sequence 980, A

```

RESULT 1
US-09-325-932A-62
: Sequence 62, Application US/09325932A
: Patent No. 6451604
: GENERAL INFORMATION
: APPLICANT: Falm, Barry
: APPLICANT: Laaham, Annette
: TITLE OF INVENTION: Compositions affectin
: TITLE OF INVENTION: death and their use i
: FILE REFERENCE: 1022
: CURRENT APPLICATION NUMBER: US/09/325,932A
: CURRENT FILING DATE: 1999-06-04
: NUMBER OF SEQ ID NOS: 206
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 62
: LENGTH: 316
: TYPE: FRT
: ORGANISM: Pinus radiata
US-09-325-932A-62

```

Query Match	72.7%;	Score 1171.5;	DB 4;	Length 316;	
Best Local Similarity	71.3%;	Pred. No. 1.4e-106;			
Matches	226;	Conservative 39;	Mismatches 51;	Indels 1;	Gaps 1
Qy	1	MATLKVDSVPAPDDAEQLRTAFEGMGNTEDLIISILAHRSABORKYIRQAYHYETGCD	60		
Db	1	MSLTIVQPLRPVADDEGQLRTAFAGMGNTNEKLIISILGHNAQRKLIROYATATYGED	60		
Qy	61	LLKTLTDELSNDFERRAILLWLTBEPGERDALLIANEATKWTSSNQYLMVEACTRISTOLLH	120		
Db	61	LLKALDEBLTNDPFRVLVWLSMDPAERDAYLANEATKWTSSNQYLMETACTRSPQOLLM	120		
Qy	121	ARQAYHARYYKKSLEEDVAHHTTGPRKLLVSLVSYRSEGEVNNATLAKQEKALVHEKK	180		
Db	121	ARQAYHARYYKKSMEEDVAHHTTGPRKLLVPLSGSYRNDGGEVNNATLAKAEKALIHKKIS	180		
Qy	181	DKHNDNDSDVRIILSTRSKAQINAFNRYQDDHGBEILSKLSEGDGDDDEFLALRSTIOCL	240		
Db	181	EKAIGHDEDLIRLIILTRSKAQVNATLHNKFNFGINDINDLKT-DPKDAFLTLLRATVTKL	239		
Qy	241	TRPELVFVDVLSAINTKGTDEGALTRIVTTRAEIDLKVIIGSEYORANSIPLKAIATYDQ	300		
Db	240	TRPEKTYREKULRIALINKGTDEGALTRVVAITRAEVDKMFISEEYORANSIPLDRAIVDQ	299		
Qy	301	RGDEYKMLVALLGEDDA 317			
Db	300	TGDYERKMLLALIGHVEA 316			





```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6593
; LENGTH: 323
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-6593
```

Query Match 29.3%; Score 472; DB 4; Length 323;

Best Local Similarity 36.3%; Pred. No. 5.6e-38; Matches 11; Conservative 63; Mismatches 130; Indels 2; Gaps 2;

```
Qy 12 APSDDAQLRTAFEGMGTNEDLIISILAHSAEQRKYIRQAYHETGYEDLLKTLDEKLSN 71
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 19 SPQVDAIQAIRGISTDEKMLISILTERSNARQIYKEYQAYKEKLDKDLGSLG 78
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Qy 72 DFERAILMTLEPGERDALLANEATKRWTSNOVLMVACTRSTOLLHARQAYHARYKK 131
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 79 HFEHLWALVTTPPAVFAKQKKSMKGAGTNEIDLILTRTRSRQMKDISQAYVTVYKK 138
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Qy 132 SLEEDVAHHTTGDPKRLVSLVTSYRGEDEVMTLAKQAKLVHEKIKKHYNDED-VI 190
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 139 SLGDDISSETSGFRKLLTLADGRDESLKVDENLAKQDQILYKAGENRMGTDEDKFT 198
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Qy 191 RIISTRSKAQINATFNRYODDHGEILKSLSEGGDDDDKFLALRSTIOCLTRPELYFVDV 250
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 199 EILCSRFPQKLTFFEDYRNISQDIYDST-KGELSGHFEDLLALVNCVRNTPAFILAER 257
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Qy 251 LRSAINKTGDEGALTRITVTRAIDLVIGEYQRNSIPLKATIKOTRGDEKMLVA 310
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 258 LHRALKGIGTDEFTLRNIMVSRSEIDLDIRTEPKKHGYGSLYSALISDTSGDYEITLLK 317
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Qy 311 LAGEDD 316
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 318 ICGGDD 323
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
```

RESULT 6  
US-09-949-016-7952

```
; Sequence 7952, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7952
; LENGTH: 334
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-7952
```

Query Match 29.3%; Score 472; DB 4; Length 334;

Best Local Similarity 36.3%; Pred. No. 5.9e-38; Matches 11; Conservative 63; Mismatches 130; Indels 2; Gaps 2;

```
Qy 12 APSDDAQLRTAFEGMGTNEDLIISILAHSAEQRKYIRQAYHETGYEDLLKTLDEKLSN 71
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 30 SPQVDAIQAIRGISTDEKMLISILTERSNARQIYKEYQAYKEKLDKDLGSLG 89
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Qy 72 DFERAILMTLEPGERDALLANEATKRWTSNOVLMVACTRSTOLLHARQAYHARYKK 131
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 90 HFEHLWALVTTPPAVFAKQKKSMKGAGTNEIDLILTRTRSRQMKDISQAYVTVYKK 149
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
```

```
Qy 132 SLEEDVAHHTTGDPKRLVSLVTSYRGEDEVMTLAKQAKLVHEKIKKHYNDED-VI 190
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 150 SLGDDISSETSGFRKLLTLADGRDESLKVDENLAKQDQILYKAGENRMGTDEDKFT 209
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Qy 191 RIISTRSKAQINATFNRYODDHGEILKSLSEGGDDDDKFLALRSTIOCLTRPELYFVDV 250
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 210 EILCSRFPQKLTFFEDYRNISQDIYDST-KGELSGHFEDLLALVNCVRNTPAFILAER 268
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Qy 251 LRSAINKTGDEGALTRITVTRAIDLVIGEYQRNSIPLKATIKOTRGDEKMLVA 310
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 269 LHRALKGIGTDEFTLRNIMVSRSEIDLDIRTEPKKHGYGSLYSALISDTSGDYEITLLK 328
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Qy 311 LAGEDD 316
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 329 ICGGDD 334
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
```

RESULT 7  
US-09-949-016-10475

```
; Sequence 10475, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10475
; LENGTH: 681
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-10475
```

Query Match 28.6%; Score 461.5; DB 4; Length 681;

Best Local Similarity 35.7%; Pred. No. 1.8e-36; Matches 112; Conservative 58; Mismatches 129; Indels 15; Gaps 4;

```
Qy 13 PSDDAQLRTAFEGMGTNEDLIISILAHSAEQRKYIRQAYHETGYEDLLKTLDEKLSND 72
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 373 PDADAALAKKMGAGTDEDTIIDITTHSNVORQIROTFSHFERDMLTDLKSEISGD 432
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Qy 73 FERRAILMTLEPGERDALLANEATKRWTSNOVLMVACTRSTOLLHARQAYHARYKKS 132
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 433 LARLILGLMMPRAHYDAKQKKAMGAGTDEKALIBILATRTVAIRAINAEVKEEDYHKS 492
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Qy 133 LSEEDVAHHTTGDPKRLVSLVTSYRGEDEVMTLAKQAKLVHEKIK-----DKHN 185
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 493 LBDALSSDTSNGHRRILISLATGHEEGE-NIDQAREDAQVAATLEADTPSSGDKTSL 551
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Qy 186 DEDVIRIISTRSKAQINATFNRY--QDDHGEILKSLSEGGDDDDKFLALRSTIOCLTR 242
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 552 ETRFMTILCTRSPHARRVFERIKMTNVDVETIKKMSGVDRDAFVAI-----VQSVN 607
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Qy 243 PELYFVDVLRSAINKTGDGALTRITVTRAIDLVIGEYQRNSIPLKATIKOTRG 302
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 608 KPLFPADKLYKSMKGAGTDEKTLTRIMVSRSEIDLINRREFLEKYDKSLHQALEBDTSG 667
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Qy 303 DYEKMLVALGEDD 316
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 668 DFLKALLACGEDD 681
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
```

RESULT 8  
US-09-949-016-10476

```

? Sequence 10476, Application US/09949016
? Patent No.6812339
? GENERAL INFORMATION:
? APPLICANT: VENTER, J. Craig et al.
? TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
? FILE REFERENCE: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
? FILE REFERENCE: CL001307
? CURRENT APPLICATION NUMBER: US/09/949,016
? PRIOR FILING DATE: 2000-04-14
? PRIOR APPLICATION NUMBER: 60/241,755
? PRIOR FILING DATE: 2000-10-20
? PRIOR APPLICATION NUMBER: 60/237,768
? PRIOR FILING DATE: 2000-10-03
? PRIOR APPLICATION NUMBER: 60/231,498
? PRIOR FILING DATE: 2000-09-08
? NUMBER OF SEQ ID NOS: 207012
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ ID NO 10476
? LENGTH: 681
? TYPE: PRT
? ORGANISM: Human
? US-09-949-016-10476

Query Match      28.6%; Score 461.5; DB 4; Length 681;
Best Local Similarity 35.7%; Pred.No.1.8e-36;
Matches 112; Conservative 58; Mismatches 129; Indels 15; Gaps 4;

QY 13 PSDAEQRTAFEGWNTNDLIIISLAHRSABQRKRYRAYHETTYGEDLLKTLDKELSD 72
Db 373 PDADKALKRKAMKGLGTDIEDTIIIDITRNSVQKQIRQTFPSHFPRDLMTDLKSIQD 432
QY 73 PERALILWTFEGERDALLIANEATYKRTYSSNOVLMEVACTRSTOLLHARQAYHARYKS 132
Db 433 LARLILGLMPRAHYDAKOLKKAMEGAGDEKALIEIILATRTVAETRAINEAKVEDYHKS 492
QY 133 LEEEDVAHHTGPRFKLYLSIVTSYVEGDENVMTLAKQAKLVHEIK-----DKHN 185
Db 493 LEDALSDTSGFRRILISLATGHRBGE-NDDQREDAOVALETLELATPPSGDKTSL 551
QY 186 DEDVIRILSTRSKAOINATFNRY---ODHGEELIKSLBEGDDDDKFLALLRSTIQCTR 242
Db 552 ETRFWTILICTRSYPHLRVFGFIMKTNYDVHETTIKEMSGVDRAFVAI----VQSVKN 607
QY 243 PELYVVDVLRSAINTGTDEGALTRIYVTRAEIDLKVCSEYQQRNSIPELKAITYKOTRG 302
Db 608 KPLFPADKLYSKMGAGTDEKTLTRIMVRSSEIDILNIRREFLEKYKSLHQAIEDTSG 667
QY 303 DYEKMLVALLGEDD 316
Db 668 DFLKALLALCGED 681

RESULT 9
US-09-949-016-7621
? Sequence 7621, Application US/09949016
? Patent No. 6812339
? GENERAL INFORMATION:
? APPLICANT: VENTER, J. Craig et al.
? TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
? FILE REFERENCE: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
? FILE REFERENCE: CL001307
? CURRENT APPLICATION NUMBER: US/09/949,016
? PRIOR FILING DATE: 2000-04-14
? PRIOR APPLICATION NUMBER: 60/241,755
? PRIOR FILING DATE: 2000-10-20
? PRIOR APPLICATION NUMBER: 60/237,768
? PRIOR FILING DATE: 2000-10-03
? PRIOR APPLICATION NUMBER: 60/231,498
? PRIOR FILING DATE: 2000-09-08
? NUMBER OF SEQ ID NOS: 207012
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ ID NO 7621
? LENGTH: 557

```

```

: TYPE: PROT
: ORGANISM: Human
US-09-949-016-7621

Query Match          28.0%; Score 451; DB 4; Length 557;
Best Local Similarity 35.7%; Pred. No. 1,4e-35;
Matches 110; Conservative 61; Mismatches 129; Indels 8; Gaps 3

QY      13  PSDDAEQRTIAPEGKGNEDLIISLHARSAPQKVIQQAHYHYGDDLLKTLDKLSND 72
Db      254  PLRAAEVLRKMKPGTDEQAIIDICLGRSKQKROQILLSPKTAQKGLIKDLKSELSGN 313

QY      73  FERAIIILMTLEGERDMLLANEATKRTMSSNOQVMEVACTSTSTQLHARQAVHARYKS 132
Db      314  FEKTLILMTKTPVLPDIYEIKRIGVGTDEACLEIILASSRNEHIRELIRAYKAEFKKT 373

QY      133  LEEEDVAHHTTGDPFKLLVSLVTSRYEGDEVNMTLAKQBAKLVEKIKDKHYNDEDVIR- 191
Db      374  LEEAIRSTSGHFGQRLILSLSQGRDESTVNDMSLAQPDAGELVYAAGENRLCTDESKFNA 433

QY      192  ILSTRKAQINATNRRYQDDHGEILKSL--EEGDDDKFLALLRSTTQCLTRPELYFV 248
Db      434  VLCSRSRAHLVAVENEYQRTMGRIEKSICREMSGDLEEGMLAV---VKCIKTPAPFA 489

QY      249  DVLASAIKGTGDBGALTRITTYTAEDIDLKTYGEGYQGRNSIPEKATKTQTRGVEYKQL 308
Db      490  ERLNKKMGAGTKDQRTILIRIMVSRSETDLDIRSEYKMYGKSLYHDISGTSGDYRKIL 549

QY      309  VALGEGDD 316
Db      550  LKICGND 557

RESULT 10
US-08-526-136-2
: Sequence 2, Application US/08526136
: Patent No. 6107089
: GENERAL INFORMATION:
: APPLICANT: Towle, Christine A. et al.
: TITLE OF INVENTION: ANNEXIN XI
: NUMBER OF SEQUENCES: 36
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Fish & Richardson
: STREET: 225 Franklin Street
: CITY: Boston
: STATE: Massachusetts
: COUNTRY: U.S.A.
: ZIP: 02110-2804
: COMPUTER READABLE FORM:
: MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
: COMPUTER: IBM PS/2 Model 502 or 55SX
: OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
: SOFTWARE: Wordperfect (Version 5.0)
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/526.136
: FILING DATE:
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US/08/214.036
: FILING DATE:
: APPLICATION NUMBER: 07/837,775
: FILING DATE: February 13, 1992
: APPLICATION NUMBER: 07/764,465
: FILING DATE: September 23, 1991
: ATTORNEY/AGENT INFORMATION:
: NAME: Clark, Paul T.
: REGISTRATION NUMBER: 30,162
: REFERENCE/DOCKET NUMBER: 00786/099001
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (617) 542-5070
: TELEFAX: (617) 542-8906
: TELEX: 200154
: INFORMATION FOR SEQ ID NO: 2:

```



Query Match	27.7%	Score 447;	DB 3;	Length 466;
Best Local Similarity	33.1%;	Pred. No. 2.7e-35;		
Matches	102;	Conservative	67;	Mismatches 121; Indels 18; Gaps 3;
QY	16	DAEOLRTAFEGWGTEDELLISILARSABORVIRQAAHETYGEDLLTKTLKEISNDPER	75	
Db	168	DAEILIRKMKGFGEIDQALIVDVVANKRSNDOROKIKAAKRTSYGKLIIIDLSELSGNNEE	227	
QY	76	ALLMTLEPGERDALLANEATRKWTSNQVIMVEVACTSTQTLARQAAYHARYKSLSE	135	
Db	228	LILALFMPPTYDAMSLRKAMQAGTQGRVLEILICTRTQNEIRREIVACYOSEFGDLEK	287	
QY	136	DVAHHTTDDFKRLVSLVTSYREEDDEVNMTLAKQEALEHKKIDKRYNDEVDYR-ILS	194	
Db	288	DIRSDTGHFRLYLIVSMQCGNRDENQSIINHQAQDARLTVQAGRGRTGTESSCFNMILIA	347	
QY	195	TRSKAQINATERNRYODDGHGEILKS-----LEEGDDDDKFLALLRSTQCLTRPELY	246	
Db	348	TRSPFQLAITHAYSRMANRDLISSVSSEFSGYVSG-----LTTIIQCALINRAAF	398	
QY	247	FVDVLRSAINKTGDEGALTRIVTTRABIDLVKIGEEYQORRNSIPLERKAITKDTRGDYEK	306	
Db	399	FAERLYYAMKAGGTDSLTVIRIVTTRSEIDLVIQIMFAQWYQKTLGTMINGDTSGDYRR	458	
QY	307	MLVALIGE 314		
Db	459	LILALVQG 466		

Query Match	27.7%;	Score 447;	DB 4;	Length 467;
Best Local Similarity	33.1%;	Pred. No. 2.7e-35;		
Matches 102;	Conservative 67;	Mismatches 121;	Indels 18;	Gaps 3

```

RESULT 15
US-09-949-016-7071
; Sequence 7071, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE. METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
;

```

; SEQ ID NO 7071  
; LENGTH: 467  
; TYPE: PRT  
; ORGANISM: Human  
US-09-949-016-7071

Query Match 27.7%; Score 447; DB 4; Length 467;  
Best Local Similarity 33.1%; Pred. No. 2,7e-35;  
Matches 102; Conservative 67; Mismatches 121; Indels 18; Gaps 3;

```
QY 16 DAQQLTAPEGKGTNEEDLISILAHSAEQKVIROAYHETYGEDLITLDEKLSNDFER 75
   |||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 169 DAEILRKAMKGFGEQAIYDVVANRSDQROKIKAFKTSYGKDLIKDLKSELGMMEE 228
   |||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
QY 76 ALLMTLEPGERDALANEATKRWTSNQVLMEVACTRTSTQLHARQAHARYKKSLEE 135
   |||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 229 LILALEFMPPTYYDAMSIRKMQAGTQERVLIELCTRTNOETREIVRCYQSEFGRDLEK 288
   |||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
QY 136 DVAHHTTGDFRKLIVSLVTSRYEGDEVNMTLAKQEAIVHEKIKDGHYNDEVDIR-ILS 194
   |||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 289 DIRSDTSGHFERLLVSMCGNNDENQSIHQMAQEDAQRLYQAGEGRLGTDSCFNWILA 348
   |||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
QY 195 TRSKAQINATFNRYQDDHGEELKS-----LEEGDDDDKFLALLRSTIQCLTRPELY 246
   |||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 349 TRSFQQLRATMEAYSRMANRDLISVSRFESGYVESG-----LKTILQCALNRPAF 399
   |||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
QY 247 FVDVLESAINKTGTDEGALTRIVTTAELDKVIIGEYQRRNSIPLKAITKDTRGDYEK 306
   |||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 400 FAERLYYAMKAGTDDSTLVIRIVTRSEIDLVOIKQFAQMYOKTLGTWIMAGDTSQDYRR 459
   |||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
QY 307 MLVALLGE 314
   |||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 460 LLLAIVGQ 467
```

Search completed: August 19, 2005, 20:28:18  
Job time : 44 secs

This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 19, 2005, 20:27:39 ; Search time 161 Seconds

(without alignments)  
771.011 Million cell updates/sec

Title: US-10-690-564-2

Perfect score: 1611

Sequence: 1 MATLKXSDSVAPSDAEOL.....KOTRDPYEKTLVALLGEDDA 317

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1759131 seqs, 391586102 residues

Total number of hits satisfying chosen parameters: 1759131

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing:

Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

Published Applications AA:\*  
1: /cgn2\_6/ptodata/2/pubppa/US07\_PUBCOMB.pep:\*  
2: /cgn2\_6/ptodata/2/pubppa/PCRT\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/2/pubppa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/2/pubppa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/2/pubppa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/ptodata/2/pubppa/PCRTUS\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/2/pubppa/US08\_NEW\_PUB.pep:\*  
8: /cgn2\_6/ptodata/2/pubppa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/2/pubppa/US09A\_PUBCOMB.pep:\*  
10: /cgn2\_6/ptodata/2/pubppa/US09B\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/2/pubppa/US09C\_PUBCOMB.pep:\*  
12: /cgn2\_6/ptodata/2/pubppa/US09C\_NEW\_PUB.pep:\*  
13: /cgn2\_6/ptodata/2/pubppa/US10A\_PUBCOMB.pep:\*  
14: /cgn2\_6/ptodata/2/pubppa/US10B\_PUBCOMB.pep:\*  
15: /cgn2\_6/ptodata/2/pubppa/US10C\_PUBCOMB.pep:\*  
16: /cgn2\_6/ptodata/2/pubppa/US10D\_PUBCOMB.pep:\*  
17: /cgn2\_6/ptodata/2/pubppa/US10E\_PUBCOMB.pep:\*  
18: /cgn2\_6/ptodata/2/pubppa/US10F\_PUBCOMB.pep:\*  
19: /cgn2\_6/ptodata/2/pubppa/US11A\_PUBCOMB.pep:\*  
20: /cgn2\_6/ptodata/2/pubppa/US11B\_PUBCOMB.pep:\*  
21: /cgn2\_6/ptodata/2/pubppa/US11C\_PUBCOMB.pep:\*  
22: /cgn2\_6/ptodata/2/pubppa/US11D\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysts of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1611	100.0	317	US-10-690-564-2	Sequence 2, Appl1
2	1185.5	73.6	316	US-10-393-840-56	Sequence 56, Appl1
3	1177.5	73.1	316	US-10-424-599-144032	Sequence 144032,
4	1177.5	73.1	333	US-10-425-114-47663	Sequence 47663, A
5	1177.5	73.1	333	US-10-425-114-55802	Sequence 55802, A
6	1171.5	72.7	316	US-10-219-220-62	Sequence 62, Appl1
7	1171.5	72.7	316	US-10-393-840-114	Sequence 114, Appl
8	1132.5	70.3	316	US-10-219-220-260	Sequence 260, App
9	1054.5	62.8	314	US-10-424-599-213299	Sequence 213299,
10	1011.5	62.7	314	US-10-767-701-45089	Sequence 45089, A
11	1009.5	62.7	339	US-10-425-114-60631	Sequence 60631, A

12	1009.5	62.7	340	US-10-425-115-186841	Sequence 186841,
13	1009.5	62.7	351	US-10-425-114-64727	Sequence 64727, A
14	1008.5	62.6	341	US-10-425-114-60513	Sequence 60513, A
15	1008.5	62.6	344	US-10-425-114-59289	Sequence 59289, A
16	1004.5	62.4	317	US-10-437-963-180790	Sequence 180790,
17	988.5	61.4	333	US-10-425-114-40052	Sequence 40052, A
18	977	60.6	315	US-10-437-963-161226	Sequence 161226,
19	975.5	60.6	314	US-10-767-701-46109	Sequence 46109, A
20	951.5	59.1	314	US-10-425-115-248356	Sequence 248356,
21	949.5	58.9	314	US-10-739-930-7820	Sequence 7820, Ap
22	919.5	57.1	245	US-10-425-114-43123	Sequence 43123, A
23	866.5	53.8	239	US-10-219-220-65	Sequence 65, Appl
24	866.5	53.8	239	US-10-393-840-54	Sequence 54, Appl
25	866.5	53.8	239	US-10-393-840-117	Sequence 117, App
26	827.5	51.4	304	US-10-424-599-231427	Sequence 231427,
27	823	51.1	312	US-10-424-599-172904	Sequence 172904,
28	799.5	49.6	313	US-10-424-599-230318	Sequence 230318,
29	786.5	48.8	313	US-10-424-599-279849	Sequence 279849,
30	786.5	48.8	320	US-10-425-114-55270	Sequence 55270, A
31	786.5	48.8	324	US-10-425-114-43962	Sequence 43962, A
32	782	48.5	314	US-10-739-930-9279	Sequence 9279, Ap
33	776	48.2	314	US-10-424-599-279847	Sequence 279847,
34	770	47.8	270	US-10-424-599-211315	Sequence 211315,
35	759	47.1	314	US-10-424-599-273745	Sequence 273745,
36	721	44.8	242	US-10-393-840-55	Sequence 55, Appl
37	691.5	42.9	289	US-10-437-963-121221	Sequence 121221,
38	635.5	39.4	241	US-10-425-114-41196	Sequence 41196, A
39	635.5	39.4	369	US-10-437-963-169438	Sequence 169438,
40	629	39.0	373	US-10-425-115-213961	Sequence 213961,
41	621.5	38.6	489	US-10-437-963-169441	Sequence 169441,
42	613.5	38.1	320	US-10-424-599-201908	Sequence 201908,
43	573	35.6	184	US-10-219-220-66	Sequence 66, Appl
44	573	35.6	184	US-10-393-840-118	Sequence 118, App
45	563	34.9	184	US-10-393-840-52	Sequence 52, Appl

#### ALIGNMENTS

RESULT 1  
US-10-690-564-2  
; Sequence 2, Appl1 US/10690564  
; Publication No. US20050089872A1  
; GENERAL INFORMATION:  
; APPLICANT: KOREA KIMHO PETROCHEMICAL CO., LTD  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING ANNEXINS FROM PLANTS  
; FILE REFERENCE:  
; CURRENT APPLICATION NUMBER: US/10/690,564  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: Kopatentlin 1.71  
; SEQ ID NO 2  
; LENGTH: 317  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
US-10-690-564-2

Query Match	100.0%	Score 1611;	DB 17;	Length 317;
Best local Similarity	100.0%	Pred. No. 1.5e-130;		
Matches 317;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1 MATLKXSDSVAPSDAEOLRTAFEGGKTNEDLIISILAHSAEQRKVIROYHYETGGD 60			
DB	1 MATLKXSDSVAPSDAEOLRTAFEGGKTNEDLIISILAHSAEQRKVIROYHYETGGD 60			
QY	61 LKTKIDKELSNDFERAILLMTLEPERDALANETKWTSSNOVLMVACTRTSTOLH 120			
DB	61 LKTKIDKELSNDFERAILLMTLEPERDALANETKWTSSNOVLMVACTRTSTOLH 120			
QY	121 AROAHYARYKSLIEDVAHHTTGDFRKLVLSTSYREGEVNNMTLAKQBAKLVHEKIX 180			
DB	121 AROAHYARYKSLIEDVAHHTTGDFRKLVLSTSYREGEVNNMTLAKQBAKLVHEKIX 180			





Query Match	73.1%;	Score 1177.5;	DB 15;	Length 333;
Best Local Similarity	71.3%;	Pred. No. 4.1e-93;		
Matches	226;	Conservative 42;	Mismatches 98;	Indels 1; Gaps 1;
Qy	1	NATLKVSDBVPASDDAEQLRTAFEEQMGNEDEIISILAHRSASEOKVIRQAVHERYGED	60	
Db	18	NATLKVPQPLPPVAVDDCEQIRKAFSGMGNEELIVSILAHRNAQRKILREYIAQTGED	77	
Qy	61	LKLTLDKELSNDFERRAILMTLEPGERDILLANEATKRWTSNQVLMENAVACTSTQTLH	120	
Db	76	LKALDKELTDSFPERLVHMLTLDASARDAFLANEATKRWTSNQVLEIACHTSSQELFA	137	
Qy	121	ARQAVHARKKSLSEEVAAHTTGDFFKLVLVLTSTRYBEGDEVNMTLAKOAVLVHEKIK	180	
Db	138	ARKAVHVLKKSLEEDVAHHTTGDFFKLLPLVSSRYREGDEVNLTAKTEAKTLHHSKS	197	
Qy	181	DGYNVEDVYIRILSTRSKAQINATFPRYODDHEELIKSLSEEDDDDKFLALRSTIQCL	240	
Db	198	NKAYVDDDIIRILATRSRAQINNTLHMYDAFQDINKDL-KDDPDEEFLSRATVVKCL	256	
Qy	241	TRPELVFVDVLSAINKTGTDEGALFRIYTTRAEIDLKVLGEEYQRRNSIPLKAIYTKOT	300	
Db	257	IRPEKVFPEKVVALLANKGTDEGALRRVVATRAEVDVLKNIADYYQRRSSVPLERATVKT	316	
Qy	301	RODYERKMLVALLGEDDA	317	
Db	317	TGDYERKMLVALLGHDDA	333	

```

RESULT 5
US-10-425-114-55802
! Sequence 55802, Application US/10425114
! Publication No. US2004003488A1
! GENERAL INFORMATION:
! APPLICANT: Liu, Jindong
! APPLICANT: Zhou, Yihua
! APPLICANT: Kovalic, David K.
! APPLICANT: Screen, Steven E.
! APPLICANT: Tabaska, Jack E.
! APPLICANT: Cao, Yongwei
! TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
! TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
! FILE REFERENCE: 38-21(53313)B
! CURRENT APPLICATION NUMBER: US/10/425,114
! CURRENT FILING DATE: 2003-04-28
! NUMBER OF SEQ ID NOS: 73128
! SEQ ID NO 55802
! LENGTH: 333
! TYPE: PRT
! ORGANISM: Glycine max
! FEATURE:
! OTHER INFORMATION: Clone ID: UC-GMOPIC033B01_FL1.pep
US-10-425-114-55802

Query Match      73.1%; Score 1177.5; DB 15; Length 333;
Best Local Similarity 71.3%; Pred. No. 4.1e-93;
Matches 226; Conservative 42; Mismatches 48; Indels 1; Gaps 1;

Oy      1      MATLVSDSVPAQDDABQLRTAFEGWGNEDLITISILHRSASEGRKVIROAHETNYGED 60
Db      18      MATLVKPPQLPFPVADDCBQLKRAFGMGNGNEELIVSILHHRNAQOKLIRETYAQTGYED 77
Oy      61      LKLTLDKELISNDFEAPILILMTLEPGERDALLANEAATKMTSSNOVIMEVACTRTSTOLLH 120
Db      78      LKALADKEITSPFERLVHLMTLDSAEERDAFLANEAATKMTSSNOVIMEVACTRTSSQLPA 137
Oy      121     AROAHYARKKSKLEEDVAHHTTGDPFKLIVSYTVRYGDEFNMTLAKQOAKLVHEKIK 180
Db      138     ARKATHVLKSKLEEDVAHHTTGDPKRLKILPLVSSRYRYGDEVNLTALATKAKLHEKIS 197
Oy      181     DKHYNDEVIRILSTRSKAQINATFNRRYODDHGEETLKSLEBGDDDKFLALIRSTIOCL 240
Db      198     NKAYNDDDFIRILARRSKQINATLNHYDAGODINKOL-KADPQDEFISLIRATVKKCL 256

```

[illegible]

```

RESULT 6
US-10-219-220-62
; Sequence 62, Application US/10219220
; Publication No. US20030082724A1
; GENERAL INFORMATION:
; APPLICANT: Flinn, Barry
; APPLICANT: Lasham, Annette
; TITLE OF INVENTION: Compositions affecting programmed cell
; TITLE OF INVENTION: death and their use in the modification of plant development
; FILE REFERENCE: 11000.1022c1
; CURRENT APPLICATION NUMBER: US/10/219,220
; CURRENT FILING DATE: 2002-08-14
; PRIOR APPLICATION NUMBER: U.S. No. US20030082724A1 09/325,932
; PRIOR FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 290
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 62
; LENGTH: 316
; TYPE: PRT
; ORGANISM: Pinus radiata
US-10-219-220-62

```

[illegible]

```

PRIORITY FILING DATE: 2000-08-10
PRIORITY APPLICATION NUMBER: US 09/170,862
PRIORITY FILING DATE: 1998-10-13
PRIORITY APPLICATION NUMBER: US 60/148,426
PRIORITY FILING DATE: 1999-08-11
PRIORITY APPLICATION NUMBER: PCT NZ/99/00169
PRIORITY FILING DATE: 1999-10-08
NUMBER OF SEQ ID NOS: 956
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 114
LENGTH: 316
TYPE: PRT
ORGANISM: Pinus radiata
US-10-393-840-114

```

Query Match	72.7%	Score	1171.5	DB	15	Length	316
Best Local Similarity	71.3%	Pred. No.	1.3e-92				
Matches	226	Conservative	39	Mismatches	51	Indels	1
						Gaps	1

QY	Db
1	1
MATLKVS	MSLT
VDSPAP	VTVP
DDAEQ	PDAD
QRTAF	DDCE
FGWGT	QRTAF
NEDELT	FGWGT
LTISL	NEDELT
ARSAE	LTISL
QKVRQ	ARSAE
AYHET	QKVRQ
YETGD	AYHET
60	YETGD
1	1
MSLT	MSLT
VTVP	VTVP
PDAD	PDAD
DDCE	DDCE
QRTAF	QRTAF
FGWGT	FGWGT
NEDELT	NEDELT
LTISL	LTISL
GRNNA	GRNNA
QRKL	QRKL
IQYAT	IQYAT
ETGD	ETGD
60	ETGD

QY 61 LKRLDKELNSDFERRAILMTLEPSPERDALLANETKRWSSNQVIMEVACTRSTQOLH 120

Db 61 LKRLDKRELNTDFERLVLVMSLDPSPERDALLANETKRWSSNQVIMEIACTRSPQOLH 120

QY 122 ARQGVHARKKSLFEDVAHHTTGDPRKLVSJMSYRYEDEDVNNTLAKQAKLYVHEKIK 180  
Db 121 ARQGVHARKKSMEDVAHHTTGDPRKLVLPGSSYRNDGDEVNNTLAKAEKILHEKIS 180

QY	181 DKHNDEDEVIRILSRKSKAQNATFNRRQODHGEELKSLSEGGDDDDKFLALRLTIQL	240
Db	181 EKAYGHEDLRLRLARRSKAQVNATLNHYKKFEGNDINKDLKT-DPKDAFLTLLRATVKCL	239

QY TRPELYFVNLRSAINKTGDEGALTRIVTTAAEIDKVGEEYORRNSPLEKAITTKDT 300  
241 |||||  
240 TREPKTFEKYLRLAINKGTDEGALTIRVAATRAEDMKFTSEYYORRNSLPDLRAIVKDT 299

QY	302	RGDYEKMLVALLGEDDA	317
	:   :		
	:   :		
Db	300	TGDEYEKMLLNLIGHVEA	316

RESULT 8  
US-10-219-220-260

Publication No. US2003008272A1  
GENERAL INFORMATION:  
APPLICANT: Film, Barry

```

1  TITLE OF INVENTION:  Compositions affecting programmed cell
2  TITLE OF INVENTION:  death and their use in the modification of plant develo
3  TITLE OF INVENTION:
4  FILE REFERENCE:  11000.1022c1  76/12,731,000

```

PRIOR FILING DATE: 2002-08-14  
CURRENT FILING DATE: 2002-08-14  
PRIOR APPLICATION NUMBER: U.S. NO. US2003008272A1 09/325,932  
PRIOR FILING DATE: 1999-06-04

```

; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 260
; LENGTH: 316
; -----

```

US-10-219-220-260  
ORGANISM: Eucalyptus grandis  
US-10-219-220-260

	Matches	218;	Conservative	48;	Mismatches	50;	Indels	1;	Gaps	1
--	---------	------	--------------	-----	------------	-----	--------	----	------	---

Db 1 M A T T A V P S V P S P A E D A E Q L Q K A F A G M G T N E D L I S I L A H R N N A Q R K V I R Q T Y A E I Y G E D 60

QY	61	LUKTLDKLSDNDFEALILLMTLEBPBRALLANEXTRXWTSNQVMEVACRSTQULL	120
Db	61	LLKALDKLSSDFEESVILLMTLDPERAPAFJSENEATKLTSSNNWLMELACRSSWELPM	120
QY	121	ARQAVHARYKKSLEEDVAHHTTGDFPKULVSVTSRYREGDEVNMTLAKQZAKLVHEKIK	180
Db	121	VRQAVHARYKKSLEEDVAVHTTTGDFPKULLVPLASFPRREGDEVNMTLAKRSEKILHEKIKH	180
QY	181	DKHYNDEVITILSTRSAQAINATFNKRYQDDHGEELKLSLEGGDDDKFLALLRSTOCL	240
Db	181	EKANVHBEILIRIVTTRSAQONATILYNNNEGMAINDL-KADPNDEFLKILRSIAIKCL	239
QY	241	TRPELVFVDVRSAINKTGDEGALLTRIVTTRAEIDLKIGEYORRNSIPELEAIKTQD	300
Db	240	TYPEKVFVKVRLAINKLTGDEBMALTRVVTTRAEVDQRIKQYHKRNSVPLDRAIANDT	299
QY	301	RGDYEKMLVALLGEDDA	317
Db	300	SGDYKRMILALVGHEDA	316

RESULT 9  
US-10-424-599-213299

Publication NO. US20040031072A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa Thomas J

APPLICANT: Zhou yihua  
 APPLICANT: Cao Yongwei  
 TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated

FILE REFERENCE: 38-21 (53223) B  
CURRENT APPLICATION NUMBER: US/10/424,599  
CURRENT FILING DATE: 2003-04-28

```

; SEQ ID NO 213299
;
; LENGTH: 319
;
; TYPE: PRT

```

```

; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(319)

```

```

; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_34633C.1.pep
US-10-424-599-213299

```

Query Match	65.5%	Score	1054.5	DB	15	Length	319
Best Local Similarity	66.8%	Pred. No.	1.6e-82				
Matches 213; Conservative	45;	Mismatches	58;	Indels	3;	Gaps	3

OY 1 MATLKVSDDPAPASDAAEOLRTAEQWGTNEDLITSLIARSABEQKVIQOAYHETTYGD 60  
| | | | : : : : | | | | : : : : | | | | : : : :  
Db 1 MATLKVPAGIPSPLEDSQLRKACQGWGNEGLISILGRHNNAQRKLIEAYSTTHGED 60

QY 6 LKLTIDKEISNDFFEAAILMTLEGGERDALLANEAATKRWTFSSNOYMEVACTRISTOLH 120  
 61 LKLTIDKEISSDFERAALVWTLDPARERDAFLANEAETKMLTSSNNWYILEIATSTRSSDILK 120  
 Db

Qy 121 AQAQYARFKKSLIEDVAHTTKGDIPLKLVPLVNSIFRYEGDEXVMNTLAKSEAKLHHEKI 180

QY 180 KDKHYNDEYIRILSTRKQOINTFNRYODDGEELKSE - EGGDDDKFALILRS - TI 237  
 181 AERAYNDEELIRILSTRKQOITATLNOYINFGAINKOLKELKQVNI CNLLRAXAI 240

```

QY      238  QCLRLPELFFVDVARSAINKTGTGEGALFTIVTTAEIDIKVIGEYQRRNSIPLEKAIT 297
        : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      241  KCLYLPPEKFFAKVRLAINKGTGEGALTRVTTAAEVDLQRIAEYQRRNSIPLEBRA 300

```



```
Db      147 TRQAVHERFRKSLSEEDIAAHVTGDFRKLVLPLVSTYRYQDPEVNTLHASEAKLHEKIH 206
      QY      181 DKRYNDEVDYRILSTSKAQINATFNRYQDDHGEELKLEEDDDDKFLALRSTIOCL 240
      Db      207 HKAYSDEDEIRILITTRSKPOLATFNHYDAFGHRINKOL-KADPDDEVYRLTRAIRIRCF 265
      QY      241 TRPELFFVDVLRSAINKGTDEGALTRIVTTTRABIDLKVIGEYORNSIPLEKAITKDT 300
      Db      266 SCDDRFFEKVAKQALAGLGTDENSLTRVITTRAEVDLKLKEAYOKRNSVPLERAVAGDT 325
      QY      301 RGDYERKMLVALLGED 315
      Db      326 SGDYESMLLALLGQE 340

RESULT 13
US-10-425-114-64727
; Sequence 64727, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 64727
; LENGTH: 361
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB4605-011-H2_FLI.pep
US-10-425-114-64727

Query Match      62.7%; Score 1009.5; DB 15; Length 361;
Best Local Similarity 60.6%; Pred. No. 1.5e-78;
Matches 191; Conservative 53; Mismatches 70; Indels 1; Gaps 1;

QY      1 MATLKVSQSDVPASDDAEQRLTAPEGWTGNETDIIISILAHRSABQKVIROAHETYGED 60
      Db      48 MATLKVPATVPVADDCQDLRKAFQCGWGTNEALIIISLGHRAAQRRAIRRAYAAYGEB 107
      QY      61 LKTLKELSNDFERAIIILMTLEPGERDALANEATKRWTSNQVLMVEVACTRSTQOLH 120
      Db      108 LRSITDEISGDFERAVILMTLDPARDAVLANEARKKPKGNRLVETIACRTISAQIFA 167
      QY      121 ARQAVHARYKKSLEEDVAHHTTGDPRKLVLPLVSTYRYGDEVNMTLAKOEAQLVHEKIK 180
      Db      168 TRQAVHERFRKSLSEEDIAAHVTGDFRKLVLPLVSTYRYQDPEVNTLHASEAKLHEKIH 227
      QY      181 DKRYNDEVDYRILSTSKAQINATFNRYQDDHGEELKLEEDDDDKFLALRSTIOCL 240
      Db      228 HKAYSDEDEIRILITTRSKPOLATFNHYDAFGHRINKOL-KADPDDEVYRLTRAIRIRCF 286
      QY      241 TRPELFFVDVLRSAINKGTDEGALTRIVTTTRABIDLKVIGEYORNSIPLEKAITKDT 300
      Db      287 SCDDRFFEKVAKQALAGLGTDENSLTRVITTRAEVDLKLKEAYOKRNSVPLERAVAGDT 346
      QY      301 RGDYERKMLVALLGED 315
      Db      347 SGDYESMLLALLGQE 361

RESULT 14
US-10-425-114-60513
; Sequence 60513, Application US/10425114
```

```
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 60513
; LENGTH: 341
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3595-049-H3_FLI.pep
US-10-425-114-60513

Query Match      62.6%; Score 1008.5; DB 15; Length 341;
Best Local Similarity 60.6%; Pred. No. 1.7e-78;
Matches 191; Conservative 53; Mismatches 70; Indels 1; Gaps 1;

QY      1 MATLKVSQSDVPASDDAEQRLTAPEGWTGNETDIIISILAHRSABQKVIROAHETYGED 60
      Db      28 MATLKVPATVPVADDCQDLRKAFQCGWGTNEALIIISLGHRAAQRRAIRRAYAAYGEB 87
      QY      61 LKTLKELSNDFERAIIILMTLEPGERDALANEATKRWTSNQVLMVEVACTRSTQOLH 120
      Db      88 LRSITDEISGDFERAVILMTLDPARDAVLANEARKKPKGNRLVETIACRTISAQIFA 147
      QY      121 ARQAVHARYKKSLEEDVAHHTTGDPRKLVLPLVSTYRYGDEVNMTLAKOEAQLVHEKIK 180
      Db      148 TRQAVHERFRKSLSEEDIAAHVTGDFRKLVLPLVSTYRYQDPEVNTLHASEAKLHEKIH 207
      QY      181 DKRYNDEVDYRILSTSKAQINATFNRYQDDHGEELKLEEDDDDKFLALRSTIOCL 240
      Db      208 HKAYSDEDEIRILITTRSKPOLATFNHYDAFGHRINKOL-KADPDDEVYRLTRAIRIRCF 266
      QY      241 TRPELFFVDVLRSAINKGTDEGALTRIVTTTRABIDLKVIGEYORNSIPLEKAITKDT 300
      Db      267 SCDDRFFEKVAKQALAGLGTDENSLTRVITTRAEVDLKLKEAYOKRNSVPLERAVAGDT 326
      QY      301 RGDYERKMLVALLGED 315
      Db      327 SGDYESMLLALLGQE 341

RESULT 15
US-10-425-114-59289
; Sequence 59289, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 59289
; LENGTH: 344
; TYPE: PRT
; ORGANISM: Zea mays subsp. mexicana
; FEATURE:
```

OTHER INFORMATION: Clone ID: UC-ZMFLTEOSINTE93E06\_FLI.pep  
US-10-425-114-59289

Query Match 62.6%; Score 1008.5; DB 15; Length 344;

Best Local Similarity 60.6%; Pred. No. 1.7e-78;

Matches 191; Conservative 53; Mismatches 70; Indels 1; Gaps 1;

```
QY 1 MATLKVSDSVPAPESDDAEOLRTAFEGWGVTNEDLIISTLAHRSAREQKVIROAHYETGED 60
   ||||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 31 MATLKVPATVPVADDCDQLKAPQGMGTNEALIIILGHDAQRRAIRRAVAEAYGEE 90
   ||||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
QY 61 LKLTLDKELSNDFERRAILMTLEPGERDALANEATKRTWTSNQVLMNEVACTRTSTQLH 120
   ||||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 91 LKRSITDEISGDFERRAVIIMTLDPARDAVLANEARKKMGKRVLYEIACTRTSAQIFA 150
   ||||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
QY 121 ARQATYARARYKSLSEEDVAHHHTGDFPKLVSLVTSYRYEGDEVNMTLAKQPAKLVHEKIK 180
   ||||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 151 TRQAYHERFPRKSLSEEDIAAHVVGDFRKLVPVSVTRYDGPVWTRLAHSEAKLHEKIH 210
   ||||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
QY 181 DKHYNDEDVIRIISTRSKAQINATFNRYODHGEELKSLSEGGDDDKFLALLRSTOCL 240
   ||||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 211 HKAYSDDDEIRILITTRSKPOLATFNHYNDAFGHRINKDL-KADPDDEYLRITLRAIIRCF 269
   ||||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
QY 241 TRPELYFVDVLRSAINTGTDEGALTRIVTTTRAEIDLVIGEYQRANSIPLKKAITKDT 300
   ||||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 270 SCFDRYFEKVARQAIAAGLTDEDSLTRVITTRAEVDLKLKEAYQKRNVSFLERAVAGDT 329
   ||||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
QY 301 RGDYKMKLVALLGED 315
   ||||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 330 SGDYESMLLALGOE 344
   ||||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
```

Search completed: August 19, 2005, 20:43:26  
Job time : 163 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 19, 2005, 20:28:24 ; Search time 179 Seconds

(without alignments)  
906.866 Million cell updates/sec

Title: US-10-690-564-2

Perfect score: 1611  
Sequence: 1 MATLKVSDSVAPSPDAAEQ.....KOTRGDYKMLVALGSDDA 317

Scoring table:

BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

UniProt 03:\*  
1: uniprot\_sprot:\*  
2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1611	100.0	317	2	09SYTO
2	1585	98.4	317	2	096527
3	1549.5	96.2	314	2	039001
4	1507	93.5	317	2	067EX8
5	1202.5	74.6	316	2	082090
6	1193.5	74.1	315	2	093157
7	1173.5	72.8	316	2	022341
8	1117.5	69.4	314	2	042657
9	1112.5	69.1	314	2	09SB88
10	1109.5	68.9	316	2	09ZRU7
11	1107	68.7	315	2	093158
12	1106.5	68.7	316	2	09XEN8
13	1096.5	68.1	314	2	024131
14	1096.5	68.1	314	2	09M3H3
15	1095.5	68.0	314	2	081536
16	1093.5	67.9	316	2	069DC2
17	1087.5	67.5	314	2	024132
18	1080	67.0	315	2	081535
19	1067.5	66.3	308	2	042922
20	1067	66.2	317	2	09XER2
21	1037.5	64.4	316	2	09LX07
22	1019.5	63.3	316	2	09CSV2
23	1009.5	62.7	314	2	043863
24	1004.5	62.4	317	2	067V21
25	991.5	61.5	318	2	0681H6
26	991.5	61.5	318	2	09LX08
27	986.5	61.2	314	2	06Z6A7
28	982.5	61.0	318	2	09CSV3
29	949.5	58.9	314	2	043864
30	879	54.6	314	1	ANX4 FRAAN
31	781.5	48.5	313	2	065848

RESULT 1	ID	Q9SYTO	PRELIMINARY;	PRT;	317 AA.
AC	Q9SYTO				
DT	01-MAY-2000	(TrEMBLrel. 13, Created)			
DT	01-MAY-2000	(TrEMBLrel. 13, Last sequence update)			
DT	05-JUL-2004	(TrEMBLrel. 27, Last annotation update)			
DE	Ca2+-dependent membrane-binding protein annexin (F14D7.2 protein).				
GN	Name=AnnAt1; Synonyms=At1G35720, F14D7.2;				
OS	Arabidopsis thaliana (Mouse-ear cress);				
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;				
OC	euroside II; Brassicales; Brassicaceae; Arabidopsi.				
OX	NCBI_TaxID=3702;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RA	Clark G.B., Roux S.J.;				
RT	"Isolation and Characterization of Two Different Arabidopsis Annexin				
RT	cDNAs (Accession Nos. AF083913 and AF083914) (PCR 99-065).";				
RL	Plant Physiol. 120:340-340(1999).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RA	Nuygen M., Karlin-Neumann G., Southwick A., Lam B., Miranda M.,				
RA	Palm C.J., Bowser L., Jones T., Bann J., Carlini P., Chen H.,				
RA	Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,				
RA	Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,				
RA	Sakurai T., Satou M., Seki M., Shim P., Yamada K., Shinozaki K.,				
RA	Ecker J., Theologis A., Davis R.W.;				
RL	Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=2088475; PubMed=12093376;				
RA	Haas B.J., Volfovsky N., Town C.D., Troukhan M., Alexandrov N.,				
RA	Feldmann K.A., Flavell R.B., White O., Salzberg S.L.;				
RT	"Full-length messenger RNA sequences greatly improve genome				
RT	annotation.";				
RL	Genome Biol. 3:RESEARCH0029-RESEARCH0029(2002).				
RN	[4]				
RP	SEQUENCE FROM N.A.				
RA	Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,				
RA	Feldmann K.;				
RL	Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.				
RN	[5]				
RP	SEQUENCE FROM N.A.				
RA	Liu S.X., Chan A.C., Sakano H., Yu G., Lee J.M., Lenz C., Pham P.,				
RA	Toriumi M., Chou J., Choi E., Chung M., Gonzalez A.,				
RA	Hwang B., Liu A., Vayberg M., Alcafi H., Brooks S., Buehler B.,				
RA	Chao Q., Conn L., Conway A.B., Hansen N.F., Johnson-Hopson C.,				
RA	Khan S., Kim C., Lam B., Miranda M., Nguyen M., Palm C.J., Shim P.,				
RA	Southwick A., Davis R.W., Ecker J.R., Federspiel N.A., Theologis A.,				
RA	Yu G.;				
RL	Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.				
RN	[6]				

32	715.5	44.4	323	2	0814Z2	0814Z2	oryza sativ
33	635.5	39.4	372	2	06L4C5	06L4C5	oryza sativ
34	602.5	37.4	321	2	09ZVJ7	09ZVJ7	arabidopsis
35	597.5	37.1	321	2	09SB45	09SB45	arabidopsis
36	579.5	36.0	257	2	094CK4	094CK4	arabidopsis
37	523	32.5	316	2	09C9X3	09C9X3	arabidopsis
38	522	32.4	316	2	09CSV4	09CSV4	arabidopsis
39	512	31.8	315	2	06H450	06H450	oryza sativ
40	511.5	31.8	333	2	09ZRS3	09ZRS3	medicago sa
41	505	31.3	316	2	06FHH6	06FHH6	homo sapien
42	502	31.2	483	2	07T391	07T391	brachydanio
43	502	31.2	526	2	0804G4	0804G4	brachydanio
44	501.5	31.1	303	2	08H5B8	08H5B8	oryza sativ
45	501	31.1	315	1	ANXD_HUMAN	P27216	homo sapien



RP SEQUENCE FROM N.A.  
 RA Theologis;  
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
 RN [7]  
 RP SEQUENCE FROM N.A.  
 RA Yamada K., Dale J.M., Hsuan V.W., Onodera C.S., Quach H.L.,  
 RA Toriumi M., Wong C., Yu H.C., Yuan S., Chen H., Cheuk R.,  
 RA Jones T., Kim C.J., Nguyen M., Palm C.J., Shinn P., Southwick A.,  
 RA Tripp M.G., Wu T., Davis R.W., Ecker J.R., Theologis A.,  
 RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.  
 RN [8]  
 RP SEQUENCE FROM N.A.  
 RA Southwick A., Nguyen M., Tripp M., Palm C.J., Jones T., Wu T.,  
 RA Carninci P., Chen H., Cheuk R., Chan M.M., Chang C.H., Dale J.M.,  
 RA Deng J.M., Hayashizaki Y., Hsuan V.W., Lee J.M., Ishida J., Kamiya A.,  
 RA Kawai J., Kim C.J., Natusaka M., Onodera C.S., Quach H.L., Sakurai T.,  
 RA Satou M., Seki M., Shinn P., Tang C.C., Toriumi M., Wong C., Yu H.C.,  
 RA Yamada K., Yu G., Yuan S., Shinzaki K., Ecker J., Theologis A.,  
 RA Davis R.W.,  
 RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: Belongs to the annexin family.  
 CC EMBL; AF083913; AAD34236.1; -.  
 DR EMBL; AY072347; AAL61954.1; -.  
 DR EMBL; AY086570; AAM63633.1; -.  
 DR EMBL; AC021198; AAF79882.1; -.  
 DR EMBL; AF332435; AAG48798.1; -.  
 DR EMBL; BT003359; AAO29977.1; -.  
 DR HSSP; P93157; INO0.  
 DR HSSP; P93157; INO0.  
 DR GO; GO:0005509; F:calcium ion binding; IEA.  
 DR GO; GO:0005544; F:calcium-dependent phospholipid binding; IEA.  
 DR InterPro: IPR001464; Annexin.  
 DR InterPro: IPR009118; Annexin\_plant.  
 DR Pfam; PF00191; Annexin; 4.  
 DR PRINTS; PRO0196; ANNEXINPLANT.  
 DR PRODOM; PD000143; Annexin; 4.  
 DR SMART; SM00335; ANX; 4.  
 DR PROSITE; PS00223; ANNEXIN; 1.  
 KW Annexin; Calcium; Calcium-binding; Calcium/phospholipid-binding;  
 KW Repeat.  
 SQ SEQUENCE 317 AA; 36203 MW; 92516D630325005F CRC64;  
 Query Match 100.0%; Score 1611; DB 2; Length 317;  
 Best Local Similarity 100.0%; Pred. No. 4, 2e-94; Indels 0; Gaps 0;  
 Matches 317; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MATLKVSDSVPAAPSDDAEQRLTAPEGWTGNETDIIISILAHRSAEQRKVIRQAVHETYGED 60  
 DB 1 MATLKVSDSVPAAPSDDAEQRLTAPEGWTGNETDIIISILAHRSAEQRKVIRQAVHETYGED 60  
 QY 61 LKLTLDKELSNDFERRAILMTLEPGERDALLANEATKRTSSNOVLMEEVACTRTSTQLH 120  
 DB 61 LKLTLDKELSNDFERRAILMTLEPGERDALLANEATKRTSSNOVLMEEVACTRTSTQLH 120  
 QY 121 AROAYHARYKKSLEEDVAHHTTGDPRKLLVSLVTSYRYEGDEVNMTLAKQBAKLVHEKIK 180  
 DB 121 AROAYHARYKKSLEEDVAHHTTGDPRKLLVSLVTSYRYEGDEVNMTLAKQBAKLVHEKIK 180  
 QY 181 DKHYNDEDVIRILSTRSKAQINATFNRYODDHGEILKSLSEGGDDDKFLALRSTIOCL 240  
 DB 181 DKHYNDEDVIRILSTRSKAQINATFNRYODDHGEILKSLSEGGDDDKFLALRSTIOCL 240  
 QY 241 TRPELYFVDVLRSAINKTGTDEGALTRIVTTAEIDLKVTGGEYQRNSIPLEKAITKOT 300  
 DB 241 TRPELYFVDVLRSAINKTGTDEGALTRIVTTAEIDLKVTGGEYQRNSIPLEKAITKOT 300  
 QY 301 RGDYKMLVALGEGDDA 317  
 DB 301 RGDYKMLVALGEGDDA 317

RESULT 2  
 ID 096527 PRELIMINARY; PRT; 317 AA.  
 AC 096527;  
 DT 01-FEB-1997 (Tremblrel. 02, Created)  
 DT 01-FEB-1997 (Tremblrel. 02, Last sequence update)  
 DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)  
 DE Arabidopsis thaliana annexin-like protein.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC euroside II; Brassicales; Brassicaceae; Arabidopsie.  
 OC NCBI\_Taxid=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Whole seedling;  
 RA MEDLINE=97008169; Pubmed=8855345; DOI=10.1073/pnas.93.20.11268;  
 RX Gidrol X., Sabelli P.A., Fern Y.S., Kush A.K.,  
 RT "Annexin-like protein from Arabidopsis thaliana rescues delta oxyR  
 mutant of Escherichia coli from H2O2 stress."  
 RL Proc. Natl. Acad. Sci. U.S.A. 93:11268-11273(1996).  
 CC -1- SIMILARITY: Belongs to the annexin family.  
 CC EMBL; U28415; AAC49472.1; -.  
 DR HSSP; P93157; INO0.  
 DR GO; GO:0005509; F:calcium ion binding; IEA.  
 DR GO; GO:0005544; F:calcium-dependent phospholipid binding; IEA.  
 DR InterPro: IPR001464; Annexin.  
 DR Pfam; PF00191; Annexin; 4.  
 DR PRINTS; PRO0196; ANNEXIN.  
 DR PRODOM; PD000143; Annexin; 4.  
 DR SMART; SM00335; ANX; 4.  
 DR PROSITE; PS00223; ANNEXIN; 1.  
 KW Annexin; Calcium; Calcium-binding; Calcium/phospholipid-binding;  
 KW Repeat.  
 SQ SEQUENCE 317 AA; 36172 MW; 375E9FA536E0D384 CRC64;  
 Query Match 98.4%; Score 1585; DB 2; Length 317;  
 Best Local Similarity 98.4%; Pred. No. 1, 8e-92; Indels 0; Gaps 0;  
 Matches 312; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
 QY 1 MATLKVSDSVPAAPSDDAEQRLTAPEGWTGNETDIIISILAHRSAEQRKVIRQAVHETYGED 60  
 DB 1 MATLKVSDSVPAAPSDDAEQRLTAPEGWTGNETDIIISILAHRSAEQRKVIRQAVHETYGED 60  
 QY 61 LKLTLDKELSNDFERRAILMTLEPGERDALLANEATKRTSSNOVLMEEVACTRTSTQLH 120  
 DB 61 LKLTLDKELSNDFERRAILMTLEPGERDALLANEATKRTSSNOVLMEEVACTRTSTQLH 120  
 QY 121 AROAYHARYKKSLEEDVAHHTTGDPRKLLVSLVTSYRYEGDEVNMTLAKQBAKLVHEKIK 180  
 DB 121 AROAYHARYKKSLEEDVAHHTTGDPRKLLVSLVTSYRYEGDEVNMTLAKQBAKLVHEKIK 180  
 QY 181 DKHYNDEDVIRILSTRSKAQINATFNRYODDHGEILKSLSEGGDDDKFLALRSTIOCL 240  
 DB 181 DKHYNDEDVIRILSTRSKAQINATFNRYODDHGEILKSLSEGGDDDKFLALRSTIOCL 240  
 QY 241 TRPELYFVDVLRSAINKTGTDEGALTRIVTTAEIDLKVTGGEYQRNSIPLEKAITKOT 300  
 DB 241 TRPELYFVDVLRSAINKTGTDEGALTRIVTTAEIDLKVTGGEYQRNSIPLEKAITKOT 300  
 QY 301 RGDYKMLVALGEGDDA 317  
 DB 301 RGDYKMLVALGEGDDA 317  
 RESULT 3  
 ID 039001 PRELIMINARY; PRT; 314 AA.  
 AC 039001;  
 DT 01-NOV-1996 (Tremblrel. 01, Created)  
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)  
 DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)

```

DE Annexin (fragment).
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Schantz R., Schantz M.L., Houline G.;
RL Submitted (JUL-1996) to the EMBL/Genbank/DBJ databases.
CC -1- SIMILARITY: Belongs to the annexin family.
CC -1- SIMILARITY: Contains 4 annexin repeats.
DR EMBL; X99224; CAA67608.1; -.
DR HSSP; P93157; IN00.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0005544; F:calcium-dependent phospholipid binding; IEA.
DR InterPro; IPR009118; Annexin.
DR Pfam; PF00191; Annexin_plant.
DR PRINTS; PR00196; ANNEXIN.
DR PRODOM; PD000143; Annexin; 4.
DR SMART; SM00335; ANX; 4.
DR - Annexin; Calcium; Calcium-binding; Calcium/phospholipid-binding;
KM Repeat.
FT NON TER 1 1
SQ SEQUENCE 314 AA; 35779 MW; D817ADCAB8E4F91 CRC64;

Query Match 96.2%; Score 1549.5; DB 2; Length 314;
Best Local Similarity 97.5%; Pred. No. 3.2e-90;
Matches 307; Conservative 2; Mismatches 5; Indels 1; Gaps 1;

QY 3 TLAKVSDVPAPSDAEQLRTAFEGWGTNEELIISILAHSAEQKRVIRQAVHETGYGDL 62
DB 1 TLAKVSDVPAPSDAEQLRTAFEGWGTNEELIISILAHSAEQKRVIRQAVHETGYGDL 60
QY 63 KTLDKESLNDPERAILMTLEPERGDRDALANEATKRTSSNQVMEVACTRTSTQQLHAR 122
DB 61 KTLDKESLNDPERAILMTLEPERGDRDALANEATKRTSSNQVMEVACTRTSTQQLHAR 120
QY 123 QAVHARYKKSLEEDVAHHTTGDPRKLLVSLVTSRYRGDEVNMTLAKQEAQLVHEKIKDK 182
DB 121 QAVHARYKKSLEEDVAHHTTGDPRKLLVSLVTSRYRGDEVNMTLAKQEAQLVHEKIKDK 180
QY 183 HYADEDEVIRILSTSKAQINATFNRYODDHGEELIKSLBEGDDDKFLALRSTIOCLTR 242
DB 181 HYADEDEVIRILSTSKAQINATFNRYODDHGEELIKSLBEGDDDKFLALRSTIOCLTR 240
QY 243 PELYFVVLNSAIKKTGTDEGALTRIVTTRAEIDLKVIIGEYORRNSIPLEKAITKOTRG 302
DB 241 PELYFVVLNSAIKKTGTDEGALTRIVTTRAEIDLKVIIGEYORRNSIPLEKAITKOTRG 299
QY 303 DYERKMLVALLGEDDA 317
DB 300 DYERKMLVALLGEDDA 314

```

## RESULT 4

```

Q6TEX8 PRELIMINARY; PRT; 317 AA.
AC 06TEX8;
DT 25-OCT-2004 (TREMBlrel. 28, Created)
DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)
DT 25-OCT-2004 (TREMBlrel. 28, Last annotation update)
DE Annexin.
OS Brassica juncea (Leaf mustard) (Indian mustard);
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Brassica.
OX NCBI_TaxID=3707;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=floral bud;

```

```

RA Jami S.K., Kirti P.B.;
RT "Annexin cDNA from the floral bud of Brassica juncea.";
RL Submitted (JUL-2003) to the EMBL/Genbank/DBJ databases.
CC -1- SIMILARITY: Belongs to the annexin family.
CC -1- SIMILARITY: Contains 4 annexin repeats.
DR EMBL; AY356355; ARI0457.1; -.
DR InterPro; IPR009118; Annexin.
DR Pfam; PF00191; Annexin; 4.
DR PRINTS; PR00196; ANNEXIN.
DR PRODOM; PD000143; Annexin; 4.
DR SMART; SM00335; ANX; 4.
DR PROSITE; PS00223; ANNEXIN.1.
KM Annexin; Calcium; Calcium-binding; Calcium/phospholipid-binding;
KW Repeat.
SQ SEQUENCE 317 AA; 36220 MW; 1A4F9211AAF39A2A CRC64;

Query Match 93.5%; Score 1507; DB 2; Length 317;
Best Local Similarity 91.2%; Pred. No. 1.6e-87;
Matches 289; Conservative 22; Mismatches 6; Indels 0; Gaps 0;

QY 1 MATLKVSDVPAPSDAEQLRTAFEGWGTNEELIISILAHSAEQKRVIRQAVHETGYGDL 60
DB 1 MATLKVSDVPAPSDAEQLRTAFEGWGTNEELIISILAHSAEQKRVIRQAVHETGYGDL 60
QY 61 LKTLDKESLNDPERAILMTLEPERGDRDALANEATKRTSSNQVMEVACTRTSTQQLHAR 120
DB 61 LKTLDKESLNDPERAILMTLEPERGDRDALANEATKRTSSNQVMEVACTRTSTQQLHAR 120
QY 121 ARQAVHARYKKSLEEDVAHHTTGDPRKLLVSLVTSRYRGDEVNMTLAKQEAQLVHEKIKDK 180
DB 121 ARQAVHARYKKSLEEDVAHHTTGDPRKLLVSLVTSRYRGDEVNMTLAKQEAQLVHEKIKDK 180
QY 181 DKHYNDEVIRILSTSKAQINATFNRYODDHGEELIKSLBEGDDDKFLALRSTIOCLTR 240
DB 181 DKHYNDEVIRILSTSKAQINATFNRYODDHGEELIKSLBEGDDDKFLALRSTIOCLTR 240
QY 241 TRPELYFVDVLSAIKKTGTDEGALTRIVTTRAEIDLKVIIGEYORRNSIPLEKAITKOT 300
DB 241 TRPELYFVDVLSAIKKTGTDEGALTRIVTTRAEIDLKVIIGEYORRNSIPLEKAITKOT 300
QY 301 RGDYERKMLVALLGEDDA 317
DB 301 RGDYERKMLVALLGEDDA 317

```

## RESULT 5

```

ID 082090 PRELIMINARY; PRT; 316 AA.
AC 082090;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Fiber annexin.
OS Gossypium hirsutum (upland cotton).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Malvales; Malvaceae; Malvaceae; Gossypium.
OX NCBI_TaxID=3635;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=Texas Marker1;
RA Shin H., Brown R.M., Jr.;
RL Submitted (FEB-1997) to the EMBL/Genbank/DBJ databases.
CC -1- SIMILARITY: Belongs to the annexin family.
CC -1- SIMILARITY: Contains 4 annexin repeats.
DR EMBL; U89609; AAC33305.1; -.
DR FJR; T31428; T31428.
DR HSSP; P93157; IN00.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0005544; F:calcium-dependent phospholipid binding; IEA.
DR InterPro; IPR001464; Annexin.

```

DR	InterPro: IPRO09118; Annexin_plant.				
DR	Pfam, PF00191; ANNEXIN.4.				
DR	PRINTS: PR00196; ANNEXIN.				
DR	PRINTS: PR01814; ANNEXINPLANT.				
DR	ProDom: PD000143; Annexin; 4.				
DR	SMART: SM00335; ANX; 4.				
DR	PROSITE: PS00223; ANNEXIN; 1.				
KW	Annexin; Calcium; Calcium-binding; Calcium/phospholipid-binding;				
KW	Repeat.				
SQ	SEQUENCE 316 AA; 36002 MW; 02A4EBB3D6A86897 CRC64;				
Query Match					
	Best Local Similarity	74.6%; Score 1202.5; DB 2; Length 316;			
	Matches 230; Conservative 37; Mismatches 45; Indels 1; Gaps 1				
QY	1 MATLKVSNSVPAPSDAEQLRTAFEGNGTNEELIISIAHSAEQRKVIROAYHTTYGD	60			
Db	1 MATLVPTTVSPVSSDCQLRKAFGMGOTNEGILIDIGHRAVEORNIIRKYATTYGD	60			
QY	61 LKLTLDKLSLNDFFEAIIIMLTPEPERALLANEA TKMTSNQVLM EAVCTSTOLLH	120			
Db	61 LLKALDKLSLNDFFELVLIMALDPERDLANEA TKMTSNQVLM EACTRSANOLLH	120			
QY	121 ARQAYHAARYKKSLEEDVAHHTTGDFRKLLSVLSYSRYEGDEVNNTLAKQEKIVHEKIT	180			
Db	121 AROAHYAARYKKSLEEDVAHHTTGDFRKLLPVSYSRYEGDEVNNTLAKTEAKLHKETS	180			
QY	181 DKHAVDEVRITILSHRSKAQTNAETNRVODHGEEILSLSGDDDDKFLALBSTOCL	240			
Db	181 DKASDDDVIVIALTRSRQAQINATINHKNKEGNDINDL-KADKDEFLLALRSTVCL	239			
QY	241 TRPELYFVDVLRSAINTGTGEDGALTRIVTTPAEIDLKVIGEYRRNSIPLEKAITDT	300			
Db	240 VYPEKYFEKVALRLAINRGTDGALTRVVCRPAEVDKLINDEYGRNSVPLTRAIIVDT	299			
QY	301 RGDYFKMLVALLG 313				
Db	300 HGDYFKMLVALLG 312				
RESULT 6					
P93157	PRELIMINARY; PRT; 315 AA.				
ID	P93157				
AC	P93157;				
DT	01-MAY-1997 (TREMBLrel. 03, Created)				
DT	01-OCT-2001 (TREMBLrel. 18, Last sequence update)				
DT	01-MAR-2004 (TREMBLrel. 26, Last annotation update)				
DE	Annexin (Fragment).				
GN	Name=Anch1;				
OS	Gossypium hirsutum (Upland cotton).				
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;				
OC	eurosid II; Malvales; Malvaceae; Malvoideae; Gossypium.				
NM	NCBI_TaxID=3635;				
RZ	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=Acala SJ-2; TISSUE=Cotton fibers;				
RX	MEDLINE=97374485; PubMed=9230934;				
RA	Delmer D.P., Potlikha T.S.;				
RT	"Structures and functions of annexins in plants.";				
RL	Cell. Mol. Life Sci. 53:546-553(1997).				
CC	-1-SIMILARITY: Belongs to the annexin family.				
CC	-1-SIMILARITY: Contains 4 annexin repeats.				
CC	EMBL: U73746; AAB67993.2; -.				
DR	PIR: T10805; T10805.				
DR	PDB: INO0; X-ray; A=1-315.				
DR	GO: GO:0005509; F:calcium-ion binding; IEA.				
DR	GO: GO:0005544; F:calcium-dependent phospholipid binding; IEA.				
DR	InterPro: IPR001464; Annexin.				
DR	InterPro: IPR009118; Annexin_plant.				
DR	Pfam: PF00191; Annexin; 4.				
DR	PRINTS: PR00196; ANNEXIN.				
DR	PRINTS: PR01814; ANNEXINPLANT.				

DR	PrcDom; P0000143; Annexin; 4.
DR	SMART; SMO0335; ANX; 4.
DR	PROSITE; PS00223; ANNEXIN; 1.
KW	Annexin; Calcium; Calcium-binding; Calcium/phospholipid-binding;
KW	Repeat.
FT	NON TER
SQ	SEQUENCE 315 AA; 35823 MW; 653113C77AAAE980 CRC64;
Query Match	74.1%; Score 1193.5; DB 2; Length 315;
Best Local Similarity	73.4%; Pred. No. 1e-67;
Matches 229; Conservative 37; Mismatches 45; Indels 1; Gaps 1	
OY	2 ATKVSDVPASPDAEOLRTAPFGMGNGEDLIISILHRSAGEORKVROAVHETYGDL 61
DB	1 ATLVPPTTPVSVDCEGLKRAFGSGKNEGIIIDLGHRAEQRLIRKYATYEYGD 60
OY	62 LKTLDELSNPERAILMTLEPEGRDALLANEATKWTSSNOVLMEVACTRSTOLLHA 121
DB	61 LKALDKELSNPFERLVLTLMALDPARDALLANEA TKWTSSNOVLMEIACTRSANOQLHA 120
OY	122 ROAVHARKSLEBDVAHTTGDDPRKLIVSLVTAYRGDEVNMNTLAQEKLVHEKID 181
DB	121 ROAVHARKSKLEBDVAHTTGDPRKLLPLVSSRYGEEVNMTLATTEKLTHEKISN 180
OY	182 KHYNDDEVIRILSTRSKAQINATFNRYODDHGEILKSLEGGDDDKFLALLRSTIOCLT 241
DB	181 KAYDDDVIRILARSKAQINATLNHYNEKGNIINDL-KADPKDELALLRSTVKCLV 239
OY	242 RPELPFDVLSAINKTGTDEGALTTRYTTTRAEIDLKVIIGEYQRNSIPLEKAITKOTR 301
DB	240 YPEKFPEFVLAIINRGTDGALT RVCTRAEVDLKVIAD EYQRNSVPLTRAIVKDTH 299
OY	302 GDYERMLVALIG 313
DB	300 GDYERKLLVLAG 311
RESULT 7	
ID 022341	PRELIMINARY; PRT; 316 AA.
AC 02341:	
DT 01-JAN-1998 (TREMBLrel. 05, Created)	
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)	
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)	
DE Annexin.	
GN Name=Anx1c1;	
OS Lavatera thuringiaca.	
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;	
OC Eusteroide II; Malvales; Malvaceae; Malvoideae; Lavatera.	
OX NCBI_TaxID=61660;	
RN [1]	
RP SEQUENCE FROM N.A.	
RRC TISSUE=Leaves without petioles;	
RA Vazquez-Tello A., Uozumi T.;	
RT "Cloning and characterization of a Lavatera thuringiaca cDNA encoding	
RL an annexin whose expression is stimulated by low temperature.";	
RL Plant Physiol. 0:0-0(1997).	
RN [2]	
RP SEQUENCE FROM N.A.	
RRC TISSUE=Leaves without petioles;	
RA MEDLINE=20255875; PubMed=10795312;	
RX Bieton G., Vazquez-Tello A., Danyluk J., Sarhan F.;	
RT "Two novel intrinsinc annexins accumulate in wheat membranes in	
RT response to low temperature.";	
CC - - SIMILARITY: Belongs to the annexin family.	
CC - - SIMILARITY: Contains 4 annexin repeats.	
DR EMEL; AF006197; AAB71830.1; -.	
DR HSSP; P93157; INO0.	
DR GO; GO:0005509; F:calcium ion binding; IEA.	
DR GO; GO:0005544; F:calcium-dependent phospholipid binding; IEA.	
DR InterPro; IPR001464; Annexin.	

DR InterPro: IPR009118; Annexin\_plant.  
 DR Pfam: PF00191; Annexin. 4.  
 DR PRINTS; PR00196; ANNEXIN.  
 DR PRINTS; PRO1814; ANNEXINPLANT.  
 DR ProDom; PD000143; Annexin. 4.  
 DR SMART; SM00335; ANX; 4.  
 DR PROSITE; PS00223; ANNEXIN; 1.  
 KW Annexin; Calcium; Calcium-binding; Calcium/phospholipid-binding;  
 KW Repeat.  
 SQ SEQUENCE 316 AA; 36083 MW; E4B906686A351AA CRC64;  
 Query Match 72.8%; Score 1173.5; DB 2; Length 316;  
 Best Local Similarity 71.0%; Pred. No. 1.9e-66;  
 Matches 225; Conservative 43; Mismatches 48; Indels 1; Gaps 1;  
 QY 1 MATLKVSQSDVPAPSDAEQRLTAPEGNGTNEDEIIISILHRSAPQRKVIQAYHETYGED 60  
 DB 1 MATLTVPSTLPVSEDCQRLKAFSGMGTEDELLIINILGHNNADERSIKRAYTETHGED 60  
 QY 61 LKLTLDKELSDNFERALLMTLTPGERDALLANEATKRWTSNQVIMEVACTRSTQLH 120  
 DB 61 LKALDELSDNFERLVLTWLPDRDALLANEATKRWTSNQVIMEIACRSSDQLR 120  
 QY 121 ARQAYHARYKKSLEEDVAHHTTGDPRKLVSLVTSYRGEDEVMTLAKQEAQVHEKIK 180  
 DB 121 ARQAYHARYKKSLEEDVAHHTTGDPRKLVSLVTSYRGEDEVMTLAKTEAKILHEKIS 180  
 QY 181 DKHYNDEVIIRILSTRSKAQINATFNRYODHGEELIKSLBEGDDDDKFLALRSTIOCL 240  
 DB 181 NKAYSDDDVIRIVLATRSKQIENRNLNHYKNEYATDINKDL-KADPKDEFLLALRSTYKCL 239  
 QY 241 TRPELVFVDVLRSAINTGTDEGALTRIVTTAEIDLVKVGEEYQRNSIPIEKATTKOT 300  
 DB 240 VPEKTFEKKVRLAINRGDEGALTVSVSTRAEVLDKIADYQRRNSVPLTRAIKOT 299  
 QY 301 RGDYEKMLVALLGEDDA 317  
 DB 300 NGDYEKMLVALLAGEVEA 316  
 RESULT 8  
 ID Q42657 PRELIMINARY; PRT; 314 AA.  
 AC Q42657;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 OS Capsicum.  
 DE Capsicum annuum (Bell pepper).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;  
 OC Lamiales; Solanales; Solanaceae; Capsicum.  
 NCBI\_TaxID=4072;  
 RN NCBI\_TaxID=4072;  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Fruit;  
 RX MEDLINE=96198603; Pubmed=8925897; DOI=10.1016/0014-5793(96)00525-9;  
 RA Proust J., Houine G., Schantz M.L., Schantz R.;  
 RT "Characterization and gene expression of an annexin during fruit  
 development in Capsicum annuum.";  
 RL FEBS Lett. 383:208-212(1996).  
 CC -1- SIMILARITY: Belongs to the annexin family.  
 CC -1- SIMILARITY: Contains 4 annexin repeats.  
 DR EMBL; X93308; CAAG3710.1; -.  
 DR PIR; S66274; S66274.  
 DR PDB; 1DK5; X-ray; A/B=1-314.  
 DR GO; GO:0005509; F:calcium ion binding; IEA.  
 DR GO; GO:0005544; F:calcium-dependent phospholipid binding; IEA.  
 DR InterPro; IPR001464; Annexin.  
 DR Pfam; PF00191; Annexin. 4.  
 DR PRINTS; PR00196; ANNEXIN.  
 DR PRINTS; PRO1814; ANNEXINPLANT.

DR ProDom; PD000143; Annexin; 4.  
 DR SMART; SM00335; ANX; 4.  
 DR PROSITE; PS00223; ANNEXIN; 1.  
 KW Annexin; Calcium; Calcium-binding; Calcium/phospholipid-binding;  
 KW Repeat.  
 SQ SEQUENCE 314 AA; 35850 MW; 891A828C44A029DB CRC64;  
 Query Match 69.4%; Score 1117.5; DB 2; Length 314;  
 Best Local Similarity 67.1%; Pred. No. 6.6e-63;  
 Matches 212; Conservative 58; Mismatches 43; Indels 3; Gaps 1;  
 QY 1 MATLKVSQSDVPAPSDAEQRLTAPEGNGTNEDEIIISILHRSAPQRKVIQAYHETYGED 60  
 DB 1 MASLTVPAPVSAAECEQRLSAFKMGTEKILIIILHRTAQKRLQYTAETFGED 60  
 QY 61 LKLTLDKELSDNFERALLMTLTPGERDALLANEATKRWTSNQVIMEVACTRSTQLH 120  
 DB 61 LKELDRLELTHDEKLVLTWLPDRDALLANEATKRWTSKSFVLVELACTRSKPELV 120  
 QY 121 ARQAYHARYKKSLEEDVAHHTTGDPRKLVSLVTSYRGEDEVMTLAKQEAQVHEKIK 180  
 DB 121 ARQAYHARYKKSLEEDVAHHTTGDPRKLVSLVTSYRGEDEVMTLAKAESKILHEKIS 180  
 QY 181 DKHYNDEVIIRILSTRSKAQINATFNRYODHGEELIKSLBEGDDDDKFLALRSTIOCL 240  
 DB 181 DKAYSDDEVIRILATRSKQIENRNLNHYKNEYATDINKDL--FVALIRATIKGL 237  
 QY 241 TRPELVFVDVLRSAINTGTDEGALTRIVTTAEIDLVKVGEEYQRNSIPIEKATTKOT 300  
 DB 238 VPEHFVFEVLDAINRGTEEDHLTRVIATRAEVLDKIADYQRRNSIPGLRAIKOT 297  
 QY 301 RGDYEKMLVALLGEDD 316  
 DB 298 RGDYBSMLLALIGQEE 313  
 RESULT 9  
 ID Q9SB88 PRELIMINARY; PRT; 314 AA.  
 AC Q9SB88;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Annexin Cap32.  
 GN Name=an.1;  
 OS Capsicum annuum (Bell pepper).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;  
 OC Lamiales; Solanales; Solanaceae; Capsicum.  
 NCBI\_TaxID=4072;  
 RN NCBI\_TaxID=4072;  
 RP SEQUENCE FROM N.A.  
 RA Proust J., Schantz R.;  
 RL Submitted (DEC-1998) to the EMBL/GenBank/DDA databases.  
 CC -1- SIMILARITY: Belongs to the annexin family.  
 CC -1- SIMILARITY: Contains 4 annexin repeats.  
 DR EMBL; AJ130829; CA10210.1; -.  
 DR HSSP; Q42657; 1DK5.  
 DR GO; GO:0005509; F:calcium ion binding; IEA.  
 DR GO; GO:0005544; F:calcium-dependent phospholipid binding; IEA.  
 DR InterPro; IPR001464; Annexin.  
 DR InterPro; IPR009118; Annexin\_plant.  
 DR Pfam; PF00191; Annexin. 4.  
 DR PRINTS; PR00196; ANNEXIN.  
 DR PRINTS; PRO1814; ANNEXINPLANT.  
 DR ProDom; PD000143; Annexin; 4.  
 DR SMART; SM00335; ANX; 4.  
 DR PROSITE; PS00223; ANNEXIN; 1.  
 KW Annexin; Calcium; Calcium-binding; Calcium/phospholipid-binding;  
 KW Repeat.  
 SQ SEQUENCE 314 AA; 35858 MW; 58C698BF0A04C0DC CRC64;  
 Query Match 69.1%; Score 1112.5; DB 2; Length 314;



```
Dh 61 LKCEKELJTSDFERAVALFTLDPARDAHLANEATKXFTSSNWLIMEIACSRSSHELLNV 120
Qy 122 RQAVHARYKKSLEEDVAHHTTGDPRKLLVSLVTSYRYEGDEVNNTLAKQEAUVHEKIKD 181
Dh 121 KKAHARYKKSLEEDVAHHTTGEYRKLLVPLVSFAFYEGDEVNNTLAKSEAKILHDKISD 180
Qy 182 KHVNDDEVIRILSTSKQAINATFNRYODDGEIELKSLEGGDDDDKFLALRSTIOCLT 241
Dh 181 KHVNDDEVIRILSTSKQAINATFNRYODDGEIELKSLEGGDDDDKFLALRSTIOCLT 239
Qy 242 RPELVFVDVLRSAINKTGDEGALTRIVTTFAEIDLKVYIGGEYORRNSIPLKATIKOTR 301
Dh 240 TPEGVFVKVLRQAINKLGSDWMALTRVVTTRAEVDVRIKAYORRNSIPLKQAIADTS 299
Qy 302 GDYERKMLVALLGEDDA 317
Dh 300 GDYERKMLVALLIGADA 315

RESULT 12
ID Q9XEN8 PRELIMINARY; PRT; 316 AA.
AC Q9XEN8;
DT * 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Vacuole-associated annexin VcAB42.
GN Name=Anx3;
OS Nicotiana tabacum (Common tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC 1mildis; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4097;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bright Yellow-2;
RA Seale D.F., Randall S.K.;
RT "Sequence Analysis of a Vacuole-Associated Annexin from Tobacco
  Plant Physiol. 119:1147-1147(1999)."
CC -1- SIMILARITY: Belongs to the annexin family.
CC -1- SIMILARITY: Contains 4 annexin repeats.
DR EMBL; AF113545; AAD24540.1; -.
DR HSSP; P93157; IN00.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0005544; F:calcium-dependent phospholipid binding; IEA.
DR InterPro; IPR001464; Annexin.
DR InterPro; IPR009118; Annexin_plant.
DR Pfam; PF00191; Annexin_4.
DR PRINTS; PR00196; ANNEXIN.
DR PRINTS; PR01814; ANNEXINPLANT.
DR ProDom; PD000143; Annexin; 4.
DR SMART; SM00335; ANX; 4.
DR PROSITE; PS00223; ANNEXIN; 1.
KW Annexin; Calcium; Calcium-binding; Calcium/phospholipid-binding;
KW Repeat.
SQ SEQUENCE 316 AA; 35937 MW; 99E7973FC04C5FBE CRC64;

Query Match 68.7%; Score 1106.5; DB 2; Length 316;
Best Local Similarity 67.5%; Pred. No. 3,3e-67;
Matches 214; Conservative 43; Mismatches 59; Indels 1; Gaps 1;
```

```
Qy 181 DKHYNDEVIRILSTSKQAINATFNRYODDGEIELKSLEGGDDDDKFLALRSTIOCL 240
Dh 181 DKHYNDEVIRILSTSKQAINATFNRYODDGEIELKSLEGGDDDDKFLALRSTIOCL 239
Qy 241 TPELVFVDVLRSAINKTGDEGALTRIVTTFAEIDLKVYIGGEYORRNSIPLKATIKOTR 300
Dh 240 KTEPKHEKVLRLAINKTGDEWDLTRVVTTRAEVDVRIKAYORRNSIPLKQAIADTS 299
Qy 301 RGDYERKMLVALLGEDDA 317
Dh 300 SGGYERKMLVALLIGADA 316

RESULT 13
ID O24131 PRELIMINARY; PRT; 314 AA.
AC O24131;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE P32.1 annexin.
GN Name=X511; Synonym=an.1;
OS Nicotiana tabacum (Common tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC 1mildis; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4097;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bright Yellow 2;
RA MEDLINE=918786; PubMed=10080701; DOI=10.1023/A:1006199814795;
RA Prouet J., Houine G., Schantz M.L., Shen W.H., Schantz R.;
RT "Regulation of biosyntheis and cellular localization of Sp32 annexins
  in tobacco BY2 cells."
RL Plant Mol. Biol. 39:361-372(1999).
CC -1- SIMILARITY: Belongs to the annexin family.
CC -1- SIMILARITY: Contains 4 annexin repeats.
DR EMBL; Y14972; CA475213.1; -.
DR EMBL; Y17502; CA476769.1; -.
DR HSSP; Q42657; IDK5.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0005544; F:calcium-dependent phospholipid binding; IEA.
DR InterPro; IPR001464; Annexin.
DR InterPro; IPR009118; Annexin_plant.
DR Pfam; PF00191; Annexin_4.
DR PRINTS; PR00196; ANNEXIN.
DR PRINTS; PR01814; ANNEXINPLANT.
DR ProDom; PD000143; Annexin; 4.
DR SMART; SM00335; ANX; 4.
DR PROSITE; PS00223; ANNEXIN; 1.
KW Annexin; Calcium; Calcium-binding; Calcium/phospholipid-binding;
KW Repeat.
SQ SEQUENCE 314 AA; 35855 MW; 9DB6F1AB3392D587 CRC64;

Query Match 68.1%; Score 1096.5; DB 2; Length 314;
Best Local Similarity 65.8%; Pred. No. 1.4e-61;
Matches 208; Conservative 59; Mismatches 46; Indels 3; Gaps 1;
```

Db 181 DKAYSDEVIIRILATRSKAQINATLNHYKDEVEDILKOLEGDE--FVGLLRATIKGL 237

QY 241 TRPELFVVDVLRSAINKGTDEGALTRIVTTTAAIDLVKVGEEYORRNISPLEKATTKDT 300

Db 238 VYBEHFVEVLRDAINRGTEBDHLTRVATRAEVDMKI1ADYEQKDS1PLGRALAKOT 297

QY 301 RGDYERKMLVALLGEDD 316

Db 298 RGDYERKMLVALLGQEE 313

RESULT 14

Q9M3H3 PRELIMINARY; PRT; 314 AA.

AC Q9M3H3; ID Q9M3H3; DT 01-OCT-2000 (TrEMBLrel. 15, Created) DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update) DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update) DE Annexin p34. GN Name=an34; OS Solanum tuberosum (Potato). OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; OC Lamids; Solanales; Solanaceae; Solanum. NCBI\_TaxID=4113; OX [1] SEQUENCE FROM N.A. RA Kandia R., Feusner I., Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases. CC -1- SIMILARITY: Belongs to the annexin family. CC -1- SIMILARITY: Contains 4 annexin repeats. DR EMBL; AJ401032; CAB92956.1; -. DR HSSP; Q42657; 1DK5. DR GO; GO:0005509; F:calcium ion binding; IEA. DR GO; GO:0005544; F:calcium-dependent phospholipid binding; IEA. DR InterPro; IPR001464; Annexin. DR InterPro; IPR009118; Annexin\_plant. DR Pfam; PF00191; Annexin\_4. DR PRINTS; PR01814; ANNEXIN. DR PRINTS; PR01814; ANNEXINPLANT. DR ProDom; PD000143; Annexin\_4. DR SMART; SM00335; ANX; 4. DR PROSITE; PS00223; ANNEXIN; 1. DR Annexin; Calcium; Calcium-binding; Calcium/phospholipid-binding; KW Repeat. SQ SEQUENCE 314 AA; 35845 MW; 9E63375C356F8C36 CRC64;

Query Match 68.1%; Score 1096.5; DB 2; Length 314; Best Local Similarity 66.5%; Pred. No. 1.4e-61; Matches 210; Conservative 56; Mismatches 47; Indels 3; Gaps 1;

QY 1 MATLKVSDSVPAASDADBLRTAFEGMGNTEDLIISLAHRSABQRKVIROAYHETYGED 60

Db 1 MASLTVPAEVPVSAEDCEQLRSAPFGWGTNEKLIISLAHRNAQRKLIROTVAETFGED 60

QY 61 LKLTLDKELSNDFERAILLWTLEPERDALLANEAATKWTSSNOVMEVAACRTSTOLH 120

Db 61 LKLELDRELTHDFEKLVLITLDPSPERDAVLAKAETKWTNSFVLVEIACRTSPREVLV 120

QY 121 ARQAYHARYKKSLEEDVAHHTTGDPRKLLSVLTSTRYEGDEVNMVTLAKOEAQVHEKIK 180

Db 121 AREAYHARNKKSLEEDVAHHTTGDHRRKLLVPLVSSYRYGDEVDLRLAKAESKVLHEKIS 180

QY 181 DKHYNDEVIRILSTRSKAQINATFNRYODDGEELIKSLSEGGDDDKFLALRSTIOCL 240

Db 181 DKAYSDBDEVIRILATRSKAQINATLNHYKDEVEDILKOLE--DDEDFVALLRATIKGL 237

QY 241 TRPELFVVDVLRSAINKGTDEGALTRIVTTTAAIDLVKVGEEYORRNISPLEKATTKDT 300

Db 238 VYBEHFVEVLRDAINRGTEBDHLTRVATRAEVDLKTIANEYQRDS1PLGRALAKOT 297

QY 301 RGDYERKMLVALLGEDD 316

Db 298 GGDYENMLVALLGQEE 313

RESULT 15

081536 PRELIMINARY; PRT; 314 AA.

AC 081536; DT 01-NOV-1998 (TrEMBLrel. 08, Created) DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update) DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update) DE Annexin p34. GN Name=AN34; OS Lycopersicon esculentum (Tomato). OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; OC Lamids; Solanales; Solanaceae; Solanum. NCBI\_TaxID=4081; OX [1] SEQUENCE FROM N.A. RA MEDLINE=99074266; PubMed=9857021; DOI=10.1074/jbc.273.52.34920; RA Lim E.K., Roberts M.R., Bowles D.J., "Biochemical characterization of tomato annexin p35. Independence of calcium binding and phosphatase activities.", J. Biol. Chem. 273:34920-34925(1998). RT [2] SEQUENCE FROM N.A. RA TISSUE=Root; RA Lim E.K., Roberts M.R., Bowles D.J., Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases. CC -1- SIMILARITY: Belongs to the annexin family. CC -1- SIMILARITY: Contains 4 annexin repeats. DR EMBL; AF079232; AAC97494.1; -. DR HSSP; Q42657; 1DK5. DR GO; GO:0005509; F:calcium ion binding; IEA. DR GO; GO:0005544; F:calcium-dependent phospholipid binding; IEA. DR InterPro; IPR001464; Annexin. DR InterPro; IPR009118; Annexin\_plant. DR Pfam; PF00191; Annexin\_4. DR PRINTS; PR01814; ANNEXIN. DR PRINTS; PR01814; ANNEXINPLANT. DR ProDom; PD000143; Annexin\_4. DR SMART; SM00335; ANX; 4. DR PROSITE; PS00223; ANNEXIN; 1. DR Annexin; Calcium; Calcium-binding; Calcium/phospholipid-binding; KW Repeat. SQ SEQUENCE 314 AA; 35801 MW; 18A7A5086A60AFAF CRC64;

Query Match 68.0%; Score 1095.5; DB 2; Length 314; Best Local Similarity 66.1%; Pred. No. 1.6e-61; Matches 209; Conservative 57; Mismatches 47; Indels 3; Gaps 1;

QY 1 MATLKVSDSVPAASDADBLRTAFEGMGNTEDLIISLAHRSABQRKVIROAYHETYGED 60

Db 1 MASLTVPAEVPVSAEDCEQLRSAPFGWGTNEKLIISLAHRNAQRKLIROTVAETFGED 60

QY 61 LKLTLDKELSNDFERAILLWTLEPERDALLANEAATKWTSSNOVMEVAACRTSTOLH 120

Db 61 LKLELDRELTHDFEKLVLITLDPSPERDAVLAKAETKWTNSFVLVEIACRTSPREVLV 120

QY 121 ARQAYHARYKKSLEEDVAHHTTGDPRKLLSVLTSTRYEGDEVNMVTLAKOEAQVHEKIK 180

Db 121 AREAYHARNKKSLEEDVAHHTTGDHRRKLLVPLVSSYRYGDEVDLRLAKAESKVLHEKIS 180

QY 181 DKHYNDEVIRILSTRSKAQINATFNRYODDGEELIKSLSEGGDDDKFLALRSTIOCL 240

Db 181 DKAYSDBDEVIRILATRSKAQINATLNHYKDEVEDILKOLE--DDEDFVALLRATIKGL 237

QY 241 TRPELFVVDVLRSAINKGTDEGALTRIVTTTAAIDLVKVGEEYORRNISPLEKATTKDT 300

Db 238 VYBEHFVEVLRDAINRGTEBDHLTRVATRAEVDLKTIANEYQRDS1PLGRALAKOT 297

QY 301 RGDYERKMLVALLGEDD 316

Db 298 GGDYENMLVALGQEB 313

Search completed: August 19, 2005, 20:46:32  
Job time : 182 secs



This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 22, 2005, 07:12:51 ; Search time 4960 Seconds

(without alignments)  
3096.835 Million cell updates/sec

Title: US-10-690-564-2

Perfect score: 1611  
Sequence: 1 MATLXVSDSVPARSDAEQV.....KOTRGDYERKMLVALLGEDDA 317

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapext 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODE=frame+g2n.model -DEV=xjh  
-Q=/cg21/USPRO.spool/US10690564/runat.19082005.172307.25292/app\_query.fasta\_1.455  
-DB=GenEmbl -QFMT=faastap -SUFFIX=rge -MINMATCH=0.1 -LOOPEX=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi -LIST=45  
-DOCALLIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=pco -NOR=exc -HEARSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USRR=US10690564\_@CG21\_1.3731\_@runat.19082005.172307.25292 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -MARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.\*  
1: gb\_ha.\*  
2: gb\_hcg.\*  
3: gb\_in.\*  
4: gb\_om.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_sts.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Length	ID	DESCRIPTION
SUMMARIES					
1	1611	100.0	954	6	CQ806050 Sequence
2	1611	100.0	954	8	AF332435 Arabidops
3	1611	100.0	1139	8	BT003359 Arabidops
4	1611	100.0	1192	8	AY086570 Arabidops

5	1611	100.0	1208	8	AY072347 Arabidops
6	1611	100.0	1230	8	AF083913 Arabidops
7	1585	98.4	1159	8	ATU28415 Arabidops
8	1549.5	96.2	1206	8	ATANNEX Arabidops
9	1507	93.5	954	8	AY356355 Arabidops
10	1450	90.0	8153	8	FL14D7 Arabidops
11	1202.5	74.6	1141	8	GHU73609 Arabidops
12	1193.5	73.6	2588	6	BD236006 Arabidops
13	1185.5	72.8	1112	8	AF006197 Arabidops
14	1173.5	72.7	1293	6	BD236034 Arabidops
15	1171.5	72.7	1293	6	AR231102 Arabidops
16	1171.5	69.5	1420	8	BT012975 Arabidops
17	1119.5	69.4	1180	8	CAANN Arabidops
18	1117.5	68.9	1080	8	CAN130956 Arabidops
19	1109.5	68.7	1146	8	GHU73747 Arabidops
20	1107	68.7	1157	6	AX927134 Arabidops
21	1106.5	68.7	1197	8	AF113545 Arabidops
22	1106.5	68.7	1197	8	AF113545 Arabidops
23	1096.5	68.1	945	8	STU401032 Arabidops
24	1096.5	68.1	1196	8	NTANNES11 Arabidops
25	1095.5	68.0	1160	8	AF079232 Arabidops
26	1093.5	67.9	1086	8	AY351650 Arabidops
27	1087.5	67.5	1172	8	NTANNES71 Arabidops
28	1080	67.0	1242	8	AF079231 Arabidops
29	1067.5	66.3	1182	8	MSANN Arabidops
30	1067	66.2	954	6	CQ806274 Arabidops
31	1067	66.2	985	8	AY096577 Arabidops
32	1067	66.2	1137	8	AF083914 Arabidops
33	1067	66.2	1160	8	AY085713 Arabidops
34	1067	66.2	1230	8	AY070400 Arabidops
35	1019.5	62.7	1005	8	AY014799 Arabidops
36	1009.5	62.7	1277	8	ZMANNP23 Arabidops
37	1004.5	62.4	1274	8	AK072559 Arabidops
38	991.5	61.5	1120	8	AK175572 Arabidops
39	991.5	61.5	1120	8	AK175641 Arabidops
40	991.5	61.5	1120	8	AK175892 Arabidops
41	982.5	61.0	1003	8	AY014798 Arabidops
42	981.5	60.9	1223	8	AK101787 Arabidops
43	949.5	58.9	1257	8	ZMANNP25 Arabidops
44	879	54.6	997	8	AF188832 Arabidops
45	877	54.4	81494	8	AB019236 Arabidops

## ALIGNMENTS

RESULT 1	CQ806050	954 bp	DNA	linear	PAT 10-MAY-2004
LOCUS	CQ806050	Sequence 2461 from Patent WO2004035798.			
DEFINITION	CQ806050				
ACCESSION	CQ806050.1	GI:47111659			
VERSION					
KEYWORDS					
SOURCE	Arabidopsis thaliana (thale cress)				
ORGANISM	Arabidopsis thaliana				
REFERENCE	1 Inze, D., de Veylder, L. and Vlieghe, K. Identification of novel e2f target genes and use thereof				
AUTHORS	Patent: WO 2004035798-A 2461 29-APR-2004;				
JOURNAL	CropDesign N.V. (BE)				
FEATURES	Location/Qualifiers				
source	1..954				
ORIGIN	/organism="Arabidopsis thaliana"				
	/mol_type="unassigned DNA"				
	/db_xref="taxon:3702"				

Alignment Scores:  
Pred. No.: 8.28e-129 Length: 954  
Score: 1611.00 Matches: 317  
Percent Similarity: 100.00% Conservative: 0

Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0
US-10-690-564-2 (1-317) x C0806050 (1-954)			
QY	1	Meta1aThirleuysValSeraspSerValPro1aProSeraspAlaGluGluLeu	20
DB	1	ATGGCGACTCTTAAGGTTCTGATTCCTGCTCCTTCTGATGATGTCGACATTTG	60
QY	21	ArgThralApheglUGlyTTPGlyThraBergLusAspleuile1leSerileuAlaHis	40
DB	61	AGAACCGCTTTTAAAGGATGGGTACCAAGAGACTGTGATCAATCAACTTGGCTCAC	120
QY	41	ArgSerAlaGluGlnArgLysVal1leArgGlnAlaTyThiSegLuthrYrGlyGluAsp	60
DB	121	AGAGTCTCAACAGAGGAAAGTCATCAGGCAAGCATCACCAAACTTACGGCGAAGAC	180
QY	61	LeuLeuLysThirleuAspLysGluLeuSerAspPheGluuArgAla1leLeuLeuTyr	80
DB	181	CTTCTCAAGACTCTTGACAGAGGCTCTTAACGATTCGAGAGACTATCTTGTGTGG	240
QY	81	ThirleuGluProGlyGluArgAspAlaLeuLeuAlaAsnGluAlaThirLysArgTyrThr	100
DB	241	ACTCTGAACCCGGTACGCTGATGCTTTATTTGGCTAATGACCTCAAAAAGATGGACT	300
QY	101	SerSerAsnGlnValLeuMetGluValAlaCysThraThrSerThrGlnLeuLeuHis	120
DB	301	TCAAGCAACCAAGTTCTTATGAGAGTTGCTTGACAGGACATCAACGAGCTGCTTAC	360
QY	121	AlaArgGlnAlaTyThiSalaArgTyLysLysSerLeuGluLusAspValAlaHisHis	140
DB	361	GCTAGGCAAGCTTACCATGCTCGCTACAGAAAGCTCTTAAGAGAGCGTGTCTCACAC	420
QY	141	ThirThirGlyAspPheArgLysLeuLeuValSerLeuValThirSerYrArgTyGlyGly	160
DB	421	ACTACCGGTACTTCGAAAGCTTTGGTTCCTTGTATCCCTCAACAGCTGCGAAGGA	480
QY	161	AspGluValAlaAsnMetThirLeuAlaLysGlnGluAlaLysLeuValHisGluLysLys	180
DB	481	GATGAACTGAACATGACATTTGGCTTAAGCAAGAAAGCTTACGATGAGAAATCAG	540
QY	181	AspLysHisTyThraAspGluAspVal1leArg1leLeuSerThraArgSerLysAlaGln	200
DB	541	GACAAAGCACTACATGATGAGATGTTATTAAGATCTTGTCCCAAGAAACAAAGCTCAG	600
QY	201	1leAspAlaThirPheAsnArgTyGlnAspAspHisGlyGluGluLutLeuLysSerLeu	220
DB	601	ATCAATGCTACTTTTAAACGCTTACCAAGATGATCATGGCAGGAAATTCACAAAGCTCTT	660
QY	221	GluGluGlyAspAspAspAspLysPheLeuAlaLeuLeuArgSerThr1leGlnCysLeu	240
DB	661	GAGGAAGAGATGATGATGACAAAGTTCTTGCATTTTGAGGTCAACCATTCAGTGTCTG	720
QY	241	ThraArgProGluLeuTyThraPheValAspValLeuArgSerAla1leAsnTyThrGlyThr	260
DB	721	ACAAGACCAAGAGCTTACTTTCGATGTTCTTCGTTCCAGCAATTCACAAACTGGAGACT	780
QY	261	AspGluGluValAlaLeuThraGlyLeuValThirThraArgAlaGluLutLeuLeuVal1le	280
DB	781	GATGAAGAGAGACTCTCAATTAATTTGACCAAGAGCTGAGATTGACTTGAAGGCTCAT	840
QY	281	GlyGluGluTyGlnArgArgAsnSer1leProLeuGluLutLysAla1leThirLysAspThr	300
DB	841	GGAGAGAGATACCAAGCGCAAGAAACAGCATTCCTTTGAGAAAGCTTATTCACAAAGACT	900
QY	301	ArgGlyAspTyGluLysMetLeuValAlaLeuLeuGlyGluAspAspAla	317
DB	901	CGTGAAGATTAACGAAAGATGCTCGTGCACCTTCGTGTAAGATGATGTGT	951
RESULT 2			
AF332435	AF332435	954 bp	mRNA linear PLN 07-MAY-2003

DEFINITION	Arabidopsis thaliana clone C00123 (e) putative Ca2+-dependent membrane-binding protein annexin (Atlg35720) mRNA, complete cds.
ACCESSION	AF332435
VERSION	AF332435.1
KEYWORDS	FLI_CDNA.
SOURCE	Arabidopsis thaliana (thale cress)
ORGANISM	Arabidopsis thaliana
REFERENCE	1 (bases 1 to 954)
AUTHORS	Yamada, K., Dale, J.M., Heuan, V.W., Onodera, C.S., Quach, H.L., Toriumi, M., Wong, C., Wu, H.C., Yu, G., Yuan, S., Chen, H., Cheuk, R., Jones, T., Kim, C.J., Nguyen, M., Palm, C.J., Shinn, P., Southwick, A., Tripp, M.G., Wu, T., Davis, R.W., Ecker, J.R. and Theologis, A.
TITLE	Arabidopsis Open Reading Frame (ORF) Clones
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 954)
AUTHORS	Yamada, K., Liu, S.X., Sakano, H., Pham, P.K., Banh, J., Ertgu, P., Lee, J.M., Toriumi, M., Yu, G., Brooks, S., Chao, Q., Chen, H., Karlin-Neumann, G., Kim, C., Lam, B., Miranda, M., Nguyen, M., Palm, C.J., Shinn, P., Southwick, A., Davis, R.W., Ecker, J.R. and Theologis, A.
TITLE	Direct Submission
JOURNAL	Submitted (28-DEC-2000) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA
AUTHORS	3 (bases 1 to 954)
REFERENCE	Yamada, K., Dale, J.M., Heuan, V.W., Onodera, C.S., Quach, H.L., Toriumi, M., Wong, C., Wu, H.C., Yu, G., Yuan, S., Chen, H., Cheuk, R., Jones, T., Kim, C.J., Nguyen, M., Palm, C.J., Shinn, P., Southwick, A., Tripp, M.G., Wu, T., Davis, R.W., Ecker, J.R. and Theologis, A.
TITLE	Direct Submission
JOURNAL	Submitted (07-MAY-2003) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA
COMMENT	Annotation based on July 2002 version of the Arabidopsis genome submitted to Genbank.
FEATURES	Location/Qualifiers
source	1..954
gene	/organism="Arabidopsis thaliana"
CDS	/mol_type="mRNA"
	/db_xref="taxon:3702"
	/chromosome="1"
	/clone="C00123 (e)"
	/ecotype="Columbia"
	/note="This clone is now in pUNI 51 vector. Previously, it was in pUNI-T3-D/V5-His-TOPO under the clone number C00123."
	1..954
	/gene="Atlg35720"
	1..954
	/gene="Atlg35720"
	/codon_start=1
	/evidence="experimental"
	/product="putative Ca2+-dependent membrane-binding protein annexin"
	/protein_id="AAG48798.1"
	/db_xref="GI:12083278"
	/translation="MATIKVDSVAPSDDAEOLRTAFEGCNGCTNEDLITSLAHSABQRVTRQAVHETVGGDLKTDKELSNPERAILMTLTFEGRDMLANEARTRKTSNOVLMVACTRTSTQLHARQAYHARVKSLEEDVAHHTTDFRLLVSLVSYREGDEVNNTLKQEKALVHEKIKQKHVDEIVIRLSRSRAQINATNRQVODDGEELIKSLSEDDDDKFLALRSTQCLTRPELVFVRLSKINKTGDEGALTRIVTTREIDLKVIEBYQRNSIPLEKAITDTRGDYERKMLVALLGDDA"
ORIGIN	
Alignment Scores:	
Pred. No.:	8,286-129
Score:	1611.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	100.00%
DB:	8
	Matches: 954
	Conservative: 317
	Mismatches: 0
	Indels: 0
	Gaps: 0

```

US-10-690-564-2 (1-317) x AF332435 (1-954)

QY 1 MetAlaThrLeuLeuValSerAspSerValProAlaProSerAspAspAlaGluGluLeu 20
Db 1 ATGGCGAGCTCTTAAGGTTTCTGATTCTCTGCTCTTCTGATGATGCTGAGCAATTG 60
QY 21 ArgThrAlaPheGluGluYTPGluYThAsnGluAspLeuIleIleSerIleLeuAlaHis 40
Db 61 AGAACCCGCTTTTGAAGATGGGATACAGAACGAGACTTGAATCATATCAATCTTGGCTCAC 120
QY 41 ArgSerAlaGluGluGlnArgValIleArgGlnAlaLeuThrArgGluYTPGluYAsp 60
Db 121 AGAAGTCTGACACGAGGAAGTCACTACGGACACATACACGAACTTACGGGAGAC 180
QY 61 LeuLeuLeuThrLeuAspLeuGluLeuSerAspPheGluYTPGluYAlaIleLeuLeuTTP 80
Db 181 CTCTCTCAAGACTCTTGAACAAGAGACTCTTACCAATTCGAGAGAGCTATCTTGTGTGG 240
QY 81 ThrLeuGluYTPGluYGluYArgAspAlaLeuLeuAlaAsnGluAlaThrLysArgTTPThr 100
Db 241 ACTCTTGAACCCGGTGAAGCGTGATGCTTTATGGCTATGAGCTACAAAAGATGGACT 300
QY 101 SerSerAsnGlnValLeuMetGluValAlaCysThrArgThrSerThrGlnLeuLeuHis 120
Db 301 TCAAGCAACCAAGTCTTATGGAAGTCTTGCAACAGACATCAACGACGCTGCTTAC 360
QY 121 AlaArgGlnAlaLeuYrHisAlaArgTyrLysLysSerLeuGluYAspValAlaHisHis 140
Db 361 GTAGAGCAAGCTTACATGCTCGCTGCTGCAAGAGTCTCTTGAAAGAGACGTTGCTCACAC 420
QY 141 ThrThrGluYAspPheArgLysLeuLeuValSerLeuValThrSerTyrArgTyrGluY 160
Db 421 ACTACCGGTGATTCGAAAGGCTTTGGTTCCTTGTACCTCATACAGTTCAGGAAGA 480
QY 161 AspGluValAlaSerMetThrLeuAlaLeuGlnGluAlaLysLeuValHisGluLysIleLys 180
Db 481 GATGAAGTGAACATGATGATGGCTTACCAAGAACTTAAGCTGCTCATGAGAAATACAG 540
QY 181 AspLysHisTyrAsnAspGluAspValIleArgIleLeuSerThrArgSerLysAlaGln 200
Db 541 GACAAGACATCAATGATGAGATGATTATTAAGATCTTGTCCACAAAGAACAAAGCTCAG 600
QY 201 IleAsnAlaThrPheAsnArgTyrGlnAspAspHisGluYTPGluYLeuYSerLeu 220
Db 601 ATCAATGCTACTTTAACCGTTTACCAAGATGATCATGCGAAGAAATCTCAAGAGCTT 660
QY 221 GluGluGluYAspAspAspAspLysPheLeuAlaLeuLeuArgSerThrIleGlnCysLeu 240
Db 661 GAGGAAGGATATGATGATGACAACTTCTTGCACTTTGAGGTCAACCATTCAGTGTG 720
QY 241 ThrArgProGluLeuYrPheValAspValLeuArgSerAlaIleAsnLysThrGluYThr 260
Db 721 ACAAGACCAAGAGCTTCTTGTGATGATGTCCTTCGTTCAACAATCAACAAACTGGAAC 780
QY 261 AspGluGluYAlaLeuThrArgIleValThrThrArgAlaGluLeuAspLeuValIle 280
Db 781 GATGAAGGACACTCATCAATTAATTTGACCAAGAACTAGATGATGATGAAGGCTCAT 840
QY 281 GluYTPGluYrGlnArgAsnSerIleProLeuGluYAspAlaIleThrLysAspThr 300
Db 841 GAGAGAGATACCAAGCGCAAGAAACGATTCCTTTGAGAAAGCTATTAACCAAGACACT 900
QY 301 ArgGluYAspTyrGluYrMetLeuValAlaLeuLeuGluYTPGluYAspAspAla 317
Db 901 CGTGAAGATTACGAGGAAGATGCTGTCGCACTTCGCGTGAAGATGATGCT 951

RESULT 3
BT003359 1139 bp mRNA linear PLN 29-JAN-2003
DEFINITION Arabidopsis thaliana Ca2+-dependent membrane-binding protein
anexin (Atig35720) mRNA, complete cds.
ACCESSION BT003359
VERSION BT003359.1 GI:28059005

```

## KEYWORDS

## SOURCE

PLI CDNA.  
Arabidopsis thaliana (thale cress)

## ORGANISM

Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.

## REFERENCE

1 (bases 1 to 1139)

## AUTHORS

Southwick,A., Nguyen,M., Tripp,M., Palm,C.J., Jones,T., Wu,T.,  
Carninci,P., Chen,H., Cheuk,R., Chan,M.M., Chang,C.H., Dale,J.M.,  
Deng,J.M., Hayashizaki,Y., Huan,V.W., Lee,J.M., Ishida,J.,  
Kamiya,A., Kawai,J., Kim,C.J., Narusaka,M., Onodera,C.S.,  
Quach,H.L., Sakurai,T., Satou,M., Seki,M., Shim,P., Tang,C.C.,  
Toroumi,M., Wong,C., Wu,H.C., Yamada,K., Yu,G., Yuan,S.,  
Shinozaki,K., Ecker,J., Theologis,A. and Davis,R.W.

## TITLE

Submitted (29-JAN-2003) DNA Sequencing and Technology Center,  
Stanford University, 855 California Avenue, Palo Alto, CA 94304,  
USA

## COMMENT

e-mail for correspondence: arab@sequence.stanford.edu

RIKEN Genomic Sciences Center (GSC) members carried out the  
collection and clustering of RAFL cDNAs (RAFL CDNA : 'RIKEN  
Arabidopsis Full-length CDNA') : Seki,M., Narusaka,M., Ishida,J.,  
Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,  
Hayashizaki,Y. and Shinozaki,K.

The Salk, Stanford, PGEC (SSP) Consortium members carried out the  
sequencing and annotation of the RAFL cDNAs : Nguyen,M., Tripp,M.,  
Southwick,A., Palm,C.J., Jones,T., Wu,T., Chen,H., Cheuk,R.,  
Chan,M.M., Chang,C.H., Dale,J.M., Deng,J.M., Huan,V.W., Lee,J.M.,  
Kim,C.J., Quach,H.L., Onodera,C.S., Shim,P., Tang,C.C.,  
Toroumi,M., Wong,C., Wu,H.C., Yamada,K., Yu,G., Yuan,S., Ecker,J.,  
Theologis,A. and Davis,R.W.

Nguyen,M., (SSP/Stanford) and Seki,M. (RIKEN GSC) contributed  
equally to this work. Shinozaki,K. (RIKEN GSC) and Davis,R.W.  
(SSP/Stanford) contributed equally to this work as PIs.

## FEATURES

## source

Location/Qualifiers

```

1..1139
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="1"
/clone="U17038"
/ecotype="Columbia"
/notes="This clone is in DENTR/SD-topo. This is a cloned
PCR product using RIKEN clone RAFL06-11-P04 (AY072347) as
a template"
1..1139
/gene="Atig35720"
1..954
/gene="Atig35720"
/codon_start=1
/product="Ca2+-dependent membrane-binding protein annexin"
/protein_id="AA029877.1"
/db_xref="GI:28059006"
/translation="MATLKVSQSVAPSDAEQLRTAFEGWGTEEDLIISILARSAR
ORRIVROAHETTYEDLTKLDEKLSNDEERAILLWTELEPGRDLALANEARWTSS
NOVINEVACTRTSTOLHARQAYHARVYHAKSEDEVAHHHTGPRKLLVLTYSRYEG
DEVNMTLAKQZAKIHEKIKDGHVNEDEYRILSTRSKQINATFRVYODDDEETLK
SIEBGDDDDKFLALRSTIQCTRPBELYVDVIRASINKTGTDGALTRIVTTBRAID
LVIGEEYGRNSIPLEKATVDTRGDEYKMLVALLGEDDA"

```

## ORIGIN

## Alignment Scores:

Pred. No.:	1 036-128	Length:	1139
Score:	1611.00	Matches:	317
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	8	Gaps:	0

US-10-690-564-2 (1-317) x BT003359 (1-1139)

```

Qy      1 MetAlaThrLeuIysValSerAspSerValProAlaProSerAspAspAlaGluGlnLeu 20
      1 ATGGCAACTCTTAAGGTTTCTGATTCGTCCCTGCTCTTCTGATGATGCTGACAAATG 60
Qy      21 ArgThrAlaPheGluGluYTrpGlyThrAsnGluAspLeuIleIleSerIleLeuAlaHis 40
      61 AGAACCGCTTTTGAAGGATGGGTACGAGACGAGACTGATTCATTCATCAATCTTGCTCAC 120
Qy      41 ArgSerAlaGluGluIleArgLysValIleArgGlnAlaSerIleGluThrTrpGlyLysAsp 60
      121 AGAAGTGTGTAACAAGAGAAAGTCATCAGGCAAGCAATACACAGAAACTTACGGCGAAGAC 180
Qy      61 LeuLeuLysThrLeuAspLysGluLeuSerAspAspPheGluLysArgAlaIleLeuLeuTrp 80
      181 CTCTCAAGACTCTTGAACAGAGACTCTTACAGATTTGACAGAGACTATCTTGTTGTGG 240
Qy      81 ThrLeuGluProGluYgluArgAspAlaLeuLeuAlaAsnGluAlaThrLysArgTrpThr 100
      241 ACTCTGAACCCGGTGGCGTGATGTCTTATGTGGCTTATGAAGCTCAAAAAGATGGACT 300
Qy      101 SerSerAspGlnAlaLeuMetGluValAlaCysThrArgThrSerThrGluLeuLeuHis 120
      301 TCAAGCAACCAAGCTTCTTATGAGAGTTGCTTGCAACAGACATCAACGACGCTGCTTAC 360
Qy      121 AlaArgGlnAlaIleThrIleAlaArgTrpLysLysSerLeuGluGluAspValAlaHisHis 140
      361 GCTAGGGAAGCTTACCATGCTGCTGCTACAGAACTCTTGAAGAGACGTTGCTTACAC 420
Qy      141 ThrThrGluAspPheArgLysLeuLeuValSerLeuValThrSerTrpArgTrpGluGly 160
      421 ACTACCGGTGACTTCAGAAAGCTTTGGTTGCTTCTTGTACTTCATACAGGTTCGAAGGA 480
Qy      161 AspGluValAsnMetThrLeuAlaLysGlnGluAlaLysLeuValHisGluLysIleLys 180
      481 GATGAAGTGAACATGACATTGGCTTACGAAAGAGCTTACGCTGCTCAATGAGAAATCAAG 540
Qy      181 AspLysHisIleTrpAsnAspGluAspValIleArgIleLeuSerThrArgSerLysAlaGln 200
      541 GACAGACACTACATGATGATGAGATGTTATGATTCCTTGCCACAGAGACAAAGCTCAG 600
Qy      201 IleAsnAlaThrPheAsnArgTrpGluAspAspHisGluGluGluIleLeuLysSerLeu 220
      601 ATCAATGCTACTTTTAAACCGTTTACCAAGATGATCATGCGACAGGAAATTTCTCAAGACT 660
Qy      221 GluGluGluYAspAspAspAspLysPheLeuAlaLeuLeuArgSerThrIleGlnCysLeu 240
      661 GAGGAGAGAGATGATGATGACAGATTCCTTGCACTTTGAGGTCAACCATTCAGTCTTG 720
Qy      241 ThrArgProGluLeuTrpPheValAspValLeuArgSerAlaIleAsnLysThrGlyThr 260
      721 ACAAGACCAAGAGCTTATCTTGTGCAGTTCCTTGTTGAGCATCAACAAACTGGAAGCT 780
Qy      261 AspGluGluYAlaLeuThrArgIleValThrThrArgAlaGluIleAspLeuLysValIle 280
      781 GATGAAGAGACACTCTCATGAATTTGTGACCAAGAGCTGATGATGACTTGAAGGTCAAT 840
Qy      281 GluYgluGluTrpGluIleArgAsnSerIleProLeuGluLysAlaIleThrLysAspThr 300
      841 GAGAGAGAGATACCAAGCGCAGAAACAGCATTCCTTTGAGAAACCTTATCAAAAGACACT 900
Qy      301 ArgGluYAspTrpGluLysMetLeuValAlaLeuLeuGluYgluLysAspAla 317
      901 CGTGGAGATTACGAGAAAGATGCTCGTGCACCTTCGGGTGAAGATGATGT 951
Db

```

```

RESULT 4
LOCUS      AY086570              1192 bp      mRNA           linear      PLN 14-APR-2003
DEFINITION Arabidopsis thaliana clone 25846 mRNA, complete sequence.
ACCESSION  AY086570
VERSION     AY086570.1  GI:21405280
KEYWORDS    FLU CDNA.
SOURCE      Arabidopsis thaliana (thale cress)

```

```

ORGANISM  Arabidopsis thaliana
           Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
           Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
           Rosids; eustosids II; Brassicales; Brassicaceae; Arabidopsids.
REFERENCE  1 (bases 1 to 1192)
AUTHORS    Haas, B.-J., Volkovskiy, N., Town, C.D., Troukhan, M., Alexandrov, N.,
           Feldmann, K.A., Flavell, R.B., White, O. and Salzberg, S.L.
TITLE      Full-length messenger RNA sequences greatly improve genome
           annotation
JOURNAL    Genome Biol. 3 (6), RESEARCH0029 (2002)
MEDLINE    22088475
PUBMED     12093376
REFERENCE  2 (bases 1 to 1192)
AUTHORS    Brover, V., Troukhan, M., Alexandrov, N., Lu, Y.-P., Flavell, R. and
           Feldmann, K.
TITLE      Full-length cDNA from Arabidopsis thaliana
JOURNAL    Unpublished
REFERENCE  3 (bases 1 to 1192)
AUTHORS    Brover, V., Troukhan, M., Alexandrov, N., Lu, Y.-P., Flavell, R. and
           Feldmann, K.
TITLE      Direct Submision
JOURNAL    Submitted (11-MAR-2002) Ceres, Inc, 3007 Malibu Canyon Road,
           Malibu, CA 90265, USA
COMMENT    This clone sequence is one of 5,000 Ceres full-length cDNAs made
           available to TIGR and Genbank. The following quality assessment of
           this set was done by comparison with known proteins: two percent of
           the clones are estimated to be 5'-truncated; less than one percent
           are 3'-truncated; approximately two percent represent alternative
           splice variants, including unspliced introns and spliced exons; one
           percent may contain premature stop codons; five percent may have
           frame shifts in a coding region. A sequence is considered to be
           5'-truncated if it lacks the translation initiation start (ATG). A
           sequence is considered to be 3'-truncated if it lacks the
           C-terminal end of the encoded protein. Please note that these cDNA
           sequences are derived from the Ws or Laer ecotypes and therefore
           may contain polymorphisms when compared to sequences from Col-0.
           Geneset carried out the library production and sequencing of the
           full-length clones. Ceres, Inc. carried out the clustering of the
           5' sequences, selection of clones, and sequence assembly.
FEATURES   source
            location/Qualifiers
            1..1192
               /organism="Arabidopsis thaliana"
               /mol_type="mRNA"
               /db_xref="taxon:3702"
               /clone="25846"
               36..989
               /codon_start=1
               /product="Ca2+-dependent membrane-binding protein annexin"
               /protein_id="AA063633.1"
               /db_xref="GI:21554612"
               /translation="MATLKVSQVAPSDDAEQLTAFEGWGTNEDLIISILAHRSAS
ORXVROAVHYETGSDILKTLDKELSNFEBAIILMTPEPGRDNLINBAATKMTSS
NOYVMEAVACTRSTQILHARQAVHARVYKSLSEEDVAHHTTGDPRLLVSLTYSVYEG
DEVNNTLAKQEKLVHEKIKDHNVDDEVIRILSTRSAQINATFNRIQDDHGEITLK
LSIEGDDDDKFLALRSTIQCTREPLFVVLKRAINKTGIDEGALTRIVYTRAEID
LKVIGEEYQRNSIPLERAKITKDTGDEYKMLVALIGBDA"
CDS
            36..989
               /codon_start=1
               /product="Ca2+-dependent membrane-binding protein annexin"
               /protein_id="AA063633.1"
               /db_xref="GI:21554612"
               /translation="MATLKVSQVAPSDDAEQLTAFEGWGTNEDLIISILAHRSAS
ORXVROAVHYETGSDILKTLDKELSNFEBAIILMTPEPGRDNLINBAATKMTSS
NOYVMEAVACTRSTQILHARQAVHARVYKSLSEEDVAHHTTGDPRLLVSLTYSVYEG
DEVNNTLAKQEKLVHEKIKDHNVDDEVIRILSTRSAQINATFNRIQDDHGEITLK
LSIEGDDDDKFLALRSTIQCTREPLFVVLKRAINKTGIDEGALTRIVYTRAEID
LKVIGEEYQRNSIPLERAKITKDTGDEYKMLVALIGBDA"
ORIGIN
Alignment Scores:
Pred. No.:      1,09e-128      Length:      1192
Score:          1611.00        Matches:      317
Percent Similarity: 100.00%    Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match:    100.00%        Indels:      0
DB:             8              Gaps:      0
US-10-690-564-2 (1-317) x AY086570 (1-1192)
Qy      1 MetAlaThrLeuIysValSerAspSerValProAlaProSerAspAspAlaGluGlnLeu 20
      1 ATGGCAACTCTTAAGGTTTCTGATTCGTCCCTGCTCTTCTGATGATGCTGACAAATG 95
Db      36 ATGGCAACTCTTAAGGTTTCTGATTCGTCCCTGCTCTTCTGATGATGCTGACAAATG 95
Qy      21 ArgThrAlaPheGluGluYTrpGlyThrAsnGluAspLeuIleIleSerIleLeuAlaHis 40

```

```

Db 96 AGAACCCGTTTGAAGATGGGATCAACAGAGACTTGAATCATATCATTTGGCTCAC 155
Qy 41 ArgSerAlaGluGlnAglValIleArgGlnAlaTyrHisGluTyrGlyGluAsp 60
Db 156 AGAAGTCTGTAACAGAGAAAGTATCGAGCAACATCATCAATCAACCTACGGTGAAGAC 215
Qy 61 LeuLeuYsrThrLeuAspLysGluLeuSerAspPheGluValAlaIleLeuLeuTyr 80
Db 216 CTTCTCAAGACTCTTGACAGAGAGCTCTTACAGACTTCGAGAGAGCTATCTTGTGG 275
Qy 81 ThrLeuGluProGlyGluArgAspAlaLeuLeuAlaSerGlnAlaThrLysArgTyrThr 100
Db 276 ACTTTGAACCCGGTGAAGCTGATGCTTATATGCTATGAAGCTACAAAAGTGAAGT 335
Qy 101 SerSerArgGlnValLeuMetGluValAlaCysThrArgThrSerThrGlnLeuLeuHis 120
Db 336 TCAAGCAACCAAGCTTATAGAGTAGCTTGACAAAGACATCAACGACGCTGCTTAC 395
Qy 121 AlaArgGlnAlaTyrHisGlnArgTyrLysSerLeuGluGluAspValAlaHisHis 140
Db 396 GCGAGGCAAGCTTACCAATGCTGCTACAGAAAGTCTTGAAGAGACGTGGCTCACAC 455
Qy 141 ThrThrGlyAspPheArgLysLeuLeuValSerLeuValThrSerTyrArgTyrGlyGly 160
Db 456 ACTACCGGTGACTTCAGAAAGCTTTGGTCTCTCTTGTACCTCATACAGTACGAAGCA 515
Qy 161 AspGluValAspMetThrLeuAlaLysGlnGlnAlaLysLeuValHisGluLysIleLys 180
Db 516 GATCAAGTGAACATGACATTTGGCTTACAGCAAGCACTAAGCTGCTCATGAAATCAAG 575
Qy 181 AspLysHisTyrAspAspGluAspValIleArgIleLeuSerThrArgSerLysAlaGln 200
Db 576 GACAGACACTACCAATGATGAGATGTTATGAAATCTTGCCACAGAGCAAGCAAGCTCAG 635
Qy 201 IleAsnAlaThrPheAsnArgTyrGlnAspAspHisGlyGluGluIleLeuLysSerLeu 220
Db 636 ATCATATCTACTTTTAAACCTTACCAAGATGATCATGGCAGAGAAATTCACAGAGCTT 695
Qy 221 GluGluGluLysAspAspAspLysPheLeuAlaLeuLeuArgSerThrIleGlnCysLeu 240
Db 696 GAGGAAGAGATGATGATGACCAAGTTCCTTGCACTTTGAGTCAACCATTCAGTGGCTTG 755
Qy 241 ThrArgProGluLeuTyrPheValAspValLeuArgSerAlaIleAsnLysThrGlyThr 260
Db 756 ACAAGACCAAGCTTACTTTCGATGATGATCTTCTGTCAGCAATCAACAAACCTGAGACT 815
Qy 261 AspGluGluAlaLeuThrArgIleValIleThrThrArgAlaGluIleAspLeuLysValIle 280
Db 816 GATCAAGAGACACTCATGAAATTTGACCAAGAGCTGAGATTCCTTGAAAGTCAATT 875
Qy 281 GlyGluGluTyrGlnArgArgAsnSerIleProLeuGluLysAlaIleThrLysAspThr 300
Db 876 GGAAGAGAGATACCGACGCAAGAACAGCATTCCTTTGAGAAAGCTATTACCAAGACACT 935
Qy 301 ArgGlyAspTyrGlyLysMetLeuValAlaLeuLeuGluGlyGluAspAla 317
Db 936 CATTGAGATTACGAGAAAGATGCTGTCGCACTTCGCTGAAAGATGATGCT 986

```

```

RESULT 5
AY072347
LOCUS AY072347 1208 bp mRNA linear PLN 20-JAN-2002
DEFINITION Arabidopsis thaliana Ca2+-dependent membrane-binding protein
anexin (At1g35720) mRNA, complete cds.
ACCESSION AY072347
VERSION AY072347.1 GI:18252242
KEYWORDS FLI CDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; euroside II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 1208)

```

```

AUTHORS Nguyen,M., Karlin-Neumann,G., Southwick,A., Lam,B., Miranda,M.,
Palm,C.J., Bower,L., Jones,T., Banh,J., Carrinzi,P., Chen,H.,
Cheuk,R., Chung,M.K., Hayashizaki,Y., Ishida,J., Kamiya,A.,
Kawai,J., Kim,C., Lin,J., Liu,S.X., Narusaka,M., Pham,P.K.,
Sakano,H., Sakurai,T., Satou,M., Seki,M., Shimn,P., Yamada,K.,
Shinozaki,K., Ecker,J., Theologis,A. and Davis,R.W.
TITLE Direct Submission
JOURNAL Submitted (02-JAN-2002) DNA Sequencing and Technology Center,
Stanford University, 855 California Avenue, Palo Alto, CA 94304,
USA
COMMENT e-mail for correspondence: arab@sequence.stanford.edu
RIKEN Genomic Sciences Center (GSC) members carried out the
collection and clustering of RAFL cDNAs (RAFL cDNA: "RIKEN
Arabidopsis Full-Length cDNA"); Seki,M., Narusaka,M., Ishida,J.,
Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,
Hayashizaki,Y. and Shinozaki,K.
The Salk, Stanford, PGEC (SSP) Consortium members carried out the
sequencing and annotation of the RAFL cDNAs: Nguyen,M.,
Southwick,A., Karlin-Neumann,G., Lam,B., Miranda,M., Palm,C.J.,
Bower,L., Jones,T., Banh,J., Chen,H., Cheuk,R., Chung,M.K.,
Kim,C., Lin,J., Liu,S.X., Pham,P.K., Sakano,H., Shimn,P.,
Yamada,K., Ecker,J., Theologis,A. and Davis,R.W.
Nguyen,M. (SSP/Stanford) and Seki,M. (RIKEN GSC) contributed
equally to this work. Shinozaki,K. (RIKEN GSC) and Davis,R.W.
(SSP/Stanford) contributed equally to this work as PIs.
FEATURES
source
1..1208
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="1"
/clone="RAFL06-11-P04"
/ecotype="Columbia"
/note="This clone is in pluescript"
1..1208
/gene="At1g35720"
33..986
/gene="At1g35720"
/codon_start=1
/product="Ca2+-dependent membrane-binding protein anexin"
/protein_id="AA161954.1"
/db_xref="GI:18252243"
/translation="MATLKVSDSVAPSDAQRLRTAFEGMGTEBDLIISILHRSAS
ORVIRIQAVHETRYGEDLKTLDKELSDNERAIIILWTLPEGRDLALANEATKRTSS
NOVIMAVACTRTSTOLHARQAVHARVKKSLDEDVYHHHTGPRKLVSLVLSVYREG
DEVNMTLAKQBAKLVHEKIKDHYNDENEDYIRLISPRSKQINATPRYQDDGSEITLK
SLBEGDDDKRFLALRSTIQCLTRPELVYDVDRSHINKTGIDEGALTRIVTTBAID
LVKIGSEYGRNSIPLEKAITDTRDGYEKMLVALGEBDA"
ORIGIN
Alignment Scores:
Pred. No.: 1, 116-128 Length: 1208
Score: 1611.00 Matches: 317
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 8 Gaps: 0
US-10-690-564-2 (1-317) x AY072347 (1-1208)
Qy 1 MetAlaThrLeuLysValSerAspSerValProAlaProSerAspAspAlaGluGlnLeu 20
Db 33 ATGGCACTCTTAAGGTTTCTGATTCGTTCCTGCTCTTCGATGATGATGCTGAGCAATTG 92
Qy 21 ArgThrAlaPheGluGluTyrGlyThrAsnGluAspLeuIleIleSerIleLeuAlaHis 40
Db 93 AGAACCGCTTTGAAAGAGGGGTACGAGAGGACCTTATCATCAATCAATCTTGGCTCAC 152
Qy 41 ArgSerAlaGluGlnArgLysValIleArgGlnAlaTyrHisGluTyrGlyGluAsp 60

```

Db 153 AGAAGTGTGAAAGAGAAAGTGCATCAGGCAAGCATACCAAGAACTTACGGCGAAGAC 212

Qy 61 LeuLeuLysThrLeuAspLysGluLeuSerAspAspPheGluArgAlaIleLeuLeuTrp 80

Db 213 CTTCTCAAGACTCTTGAACAAGAGAGCTCTTCAAGAAATTTCAAGAGACTATCTTGGTGG 272

Qy 81 ThrLeuGluProGlyGluArgAspAlaLeuLeuAlaAsnGluAlaThrLysArgTPrThr 100

Db 273 ACTCTTGAACCCGGTAGCGCTGATGCTTATTTGGCTTAATGAACTTCAAAAAGATGAGACT 332

Qy 101 SerSerAsnGluValLeuLeuGluValAlaCysThrArgThrSerThrGluLeuHis 120

Db 333 TCAAGAACCAAGCTTCTTATGGAAGTGTCTTGCAACAAGACATCAACGCGAGCTTCTCAC 392

Qy 121 AlaArgGluAlaLeuThrHisAlaArgTyrLysSerLeuGluGluAspValAlaHisHis 140

Db 393 GCTAAGCAAGCTTACCAATGCTGCTCAAGAAAGCTCTTGAAGAGACGTTGCTCACAC 452

Qy 141 ThrThrGlyAspPheArgLysLeuLeuValSerLeuValThrSerTyrArgTyrGluGly 160

Db 453 ACTACCGGTGACTTTCAGAAAGCTTTTGGTTTCTCTTGTTACTTACATACAGTACGAAGA 512

Qy 161 AspGluValAspMetThrLeuAlaLysGluGluAlaLysLeuValHisGluLysIleLys 180

Db 513 GATGAAGTGAAACATGACATTGGCTTAAGCAAGAGCTTAAGCTGTCATGAGAAATCAAG 572

Qy 181 AspLysHisTyrTrpAspArgLysAspValIleArgIleLeuSerThrArgSerLysAlaGln 200

Db 573 GACAAAGCACTACATAGATGATGACAAAGTTCCTTGGAGTTCGCAACAAGCAAGCAAGCTCAG 632

Qy 201 IleAspAlaThrPheAsnArgTyrGlnAspAspHisGlyGluGluLysLeuLysSerLeu 220

Db 633 ATCAATGCTACTTTTAAACCTTACCAAGATGATCATGCGCAGAAATTTCTCAAGAGTCTT 692

Qy 221 GluGluGlyAspAspAspAspLysPheLeuAlaLeuLeuArgSerThrIleGlnCysLeu 240

Db 693 GAGGAAGAGATGATGATGACAAAGTTCCTTGGAGTTCGCAACAACATTCAGTCTTGG 752

Qy 241 ThrArgProGluLeuLysThrPheValAspValLeuArgSerAlaIleAsnLysThrGlyLys 260

Db 753 ACAAGACCAAGAGCTTTCATTTGCGAGTTCCTTGGTTCGCAACATCAACAAAGCTGGAAC 812

Qy 261 AspGluGlyAlaLeuThrArgIleValThrThrArgAlaGluIleAspLeuLysValIle 280

Db 813 GATGAAGAGCACTTCACTAGATTAATTTGATGCCAAGAGCTGAGTTGACTTGAAGGCTCAT 872

Qy 281 GlyGluGluLysThrGlnArgArgAsnSerIleProLeuGluLysAlaIleThrLysAspThr 300

Db 873 GGAGAGGAGTACCAAGCGCAGGAACAGCATTTCTTTGAGAAAGCTTATTCACAAAGACACT 932

Qy 301 ArgGlyAspTyrGluLysMetLeuValAlaLeuLeuGlyGluAspAspAla 317

Db 933 CGTGAAGATTACGAAGATGCTGTCGCACTTCTCGGTGAAGATGATGCT 983

RESULT 6

AF083913 1230 bp mRNA linear PLN 02-JUN-1999

LOCUS AF083913

DEFINITION Arabidopsis thaliana annexin (AnnAcl) mRNA, complete cds.

ACCESSION AF083913

VERSION AF083913.1

KEYWORDS GI:4959105

SOURCE Arabidopsis thaliana

ORGANISM Arabidopsis thaliana (chale cress)

REFERENCE 1 Arabidopsis thaliana (chale cress);

AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

TITLE Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

JOURNAL Fossils II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE 1 (bases 1 to 1230)

CDNAs (Accession Nos. AF083913 and AF083914) (PCR 99-065)

JOURNAL Plant Physiol. 120 (1), 340 (1999)

REFERENCE 2 (bases 1 to 1230)

AUTHORS Clark, G.B. and Roux, S.J.

Clark, G.B. and Roux, S.J.

TITLE Direct Submission

JOURNAL Submitted (12-AUG-1998) Botany, University of Texas, 24th at

Whitis, Austin, TX 78713, USA

FEATURES

source

1. .1230

/organism="Arabidopsis thaliana"

/mol\_type="mRNA"

/cuiivar="Columbia"

/db\_xref="taxon:3702"

1. .1230

/gene="AnnAcl1"

34. .987

/gene="AnnAcl1"

/note="calcium-binding protein"

/codon\_start=1

/product="annexin"

/protein\_id="AAD34236.1"

/db\_xref="GI:4959106"

/translation="MATLKVSQSDVPAPDDAEQLRTAFEGMGTNEDLIISILAHRSAS

ORKVROAVHETRGEDLKTLDKESLNFERRALIMLTPEGRDLALNVEATKRTSS

NOVMEVACTRSTQLLHARQAYHARYSKSLSEDAVHTTGDPRLLSVLTSVYEG

DEVNNTLAKQEKLVHEKIKDHRNDEVDVIRLSTRKAQINATRNRYQDDHGEELIK

SLRECDDDKFLALIRSTIQCLTRPELVFVDLRSAINKGTGDBGALTRIVTTAEID

LKVIQEEYQRRNSIPLERAKITKDTGDEKMLVALLEGDDA"

ORIGIN

Alignment Scores:

Pred. No.: 1 13e-128 Length: 1230

Score: 1611.00 Matches: 317

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 8 Gaps: 0

US-10-690-564-2 (1-317) x AF083913 (1-1230)

Qy 1 MetAlaThrLeuLysValSerAspSerValProAlaProSerAspAspAlaGluGluLeu 20

Db 34 ATGGGAGACTTAAAGTTTCTGATTCGTGCTGCTCTTGCATGATGCTGAGCAATTG 93

Qy 21 ArgThrAlaPheGluGlyTPrGlyThrAsnGluAspLeuIleIleSerIleLeuAlaHis 40

Db 94 AGAACCGCTTTGGAAGAGTGGGTACGAACGAGCACTTGATCATATCAATCTTGGCTCAC 153

Qy 41 ArgSerAlaGluGluArgLysValIleArgGlnAlaTyrHisGluThrTyrGluLysAsp 60

Db 154 AGAAGTGTGAAAGAGAAAGTGCATCAGGCAAGCATACCAAGAACTTACGGCGAAGAC 213

Qy 61 LeuLeuLysThrLeuAspLysGluLeuSerAspAspPheGluArgAlaIleLeuLeuTrp 80

Db 214 CTTCTCAAGACTCTTGAACAAGAGAGCTCTTACGATTTGAGAGAGCTATCTTGGTGG 273

Qy 81 ThrLeuGluProGlyGluArgAspAlaLeuLeuAlaAsnGluAlaThrLysArgTPrThr 100

Db 274 ACTCTTGAACCCGGTAGCGCTGATGCTTATTTGGCTTAATGAACTTCAAAAAGATGACT 333

Qy 101 SerSerAsnGluValLeuLeuGluValAlaCysThrArgThrSerThrGluLeuHis 120

Db 334 TCAAGAACCAAGCTTCTTATGGAAGTGTCTTGCAACAAGACATCAACGCGAGCTTCTCAC 393

Qy 121 AlaArgGluAlaLeuThrHisAlaArgTyrLysSerLeuGluGluAspValAlaHisHis 140

Db 394 GCTAAGCAAGCTTACATGCTGCTTACGAAGAGTCTTTGAAGAGAGAGCTTGTCTCACAC 453

Qy 141 ThrThrGlyAspPheArgLysLeuLeuValSerLeuValThrSerTyrArgTyrGluGly 160

Db 454 ACTACCGGTGACTTTCAGAAAGCTTTTGGTTTCTCTTGTTACTTACATACAGTACGAAGA 513

Qy 161 AspGluValAspMetThrLeuAlaLysGluGluAlaLysLeuValHisGluLysIleLys 180

Db 514 GATGAAGTGAAACATGACATTGGCTTAAGCAAGAGCTTAAGCTGTCATGAGAAATCAAG 573

Qy 181 AspLysHisTyrAspAspArgLysAspValIleArgIleLeuSerThrArgSerLysAlaGln 200

```

Db      574 GACAGACATTCATGATGATGAGATCTTATTCATCTTCTCCACAAAGCAAGCTCAG 633
Qy      201 TLeaanaLathrPheanaaryTyrGlnaAspaPhisgljgluLuleuLysSerLeu 220
Db      634 ATCAATGCTACTTTAAACCGTTACCAAGATGATCATGGCGAGAAATTTCTCAAGACTT 693
Qy      221 GlnGlnGlnLysAspaAspaAspaAspaAspaAspaAspaAspaAspaAspaAspa 240
Db      694 GAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 753
Qy      241 ThrArgProGluLeuTyrPheValAspValLeuArgSerAlaLeuAnlySerThrGlyThr 260
Db      754 ACAAGACCAAGACTTATCTTTGATGATCTTCTTCTTCAAGCATCAACAAAGCTGAACT 813
Qy      261 AspGlnGlnAlaLeuThrArgLeuValThrThrArgAlaGlnLysAspLeuLysValIle 280
Db      814 GATGAAAGACACACTCATTAGATTTGATGACCAAGAGCTGAGATTGACTTGAAGCTCAT 873
Qy      281 GlnGlnGlnLysTyrGlnArgArgAsnSerIleProLeuGlnLysAlaIleThrLysAspThr 300
Db      874 GGAAGAGAGATCCAGCGCCAGAGACAGCATTCCTTTGAGAAAGCTATTACCAAGACACT 933
Qy      301 ArgGlyAspTyrGlnLysMetLeuValAlaLeuLeuGlnLysAspAspAla 317
Db      934 CGTGAGATTACGAGAAAGATCTCTGCTGCACTTCTCGTGAAGATGATGCT 984

RESULT 7
ATU28415      1159 bp      mRNA      linear      PLN 21-OCT-1996
LOCUS         Arabidopsis thaliana annexin-like protein mRNA, complete cds.
DEFINITION   U28415.1 GI:1621538
VERSION      U28415.1 GI:1621538
KEYWORDS
SOURCE       Arabidopsis thaliana (chale cress)
ORGANISM     Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosids; eurosoid II; Brassicales; Brassicaceae; Arabidopsids.
1. (bases 1 to 1159)
Gidrol.X., Sabelli.P.A., Fern.Y.S. and Kush.A.K.
Annexin-like protein from Arabidopsis thaliana rescues delta oxyr
mutant of Escherichia coli from H2O2 stress
Proc. Natl. Acad. Sci. U.S.A. 93 (20), 11268-11273 (1996)

JOURNAL      MEDLINE      9708169
PUBMED       8853345
REFERENCE    2. (bases 1 to 1159)
AUTHORS      Kush.A.K.
TITLE        Direct Submission
JOURNAL      Submitted (06-JUN-1995) Anil K. Kush, Institute of Molecular & Cell
Biology, National University of Singapore, 10 Kent Ridge Crescent,
Singapore 0511, Republic of Singapore
location/Qualifiers
1. 1159
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/variety="ecotype Columbia"
/db_xref="taxon:3702"
/clone="pox5"
/tissue_type="whole seedling"
/clone_lib="lambda UniZapit cDNA library of A.K.Kush"
/dev_stage="4 week-old"
6. .959
/note="annexin-like protein"
/codon_start=1
/protein_id="AAC49472.1"
/db_xref="GI:1621538"
/translation="MATLKYSDSVPAPSDAEOLRTPFSGMGTEMDLIISILAHSAE
QRVIRQAHETYGKDLTKTLDELSENDFERALLWTLGEGEDALLANEATRKWTSS
NOYLMEVATKSTLQHAQVHAARKSLBEDVAHHTGDRKLIVLSRYEG
DEVNMLAKQEAHTLVEHKIKDKAHNDEVDVRISTRSKQINATPNRYODHGEELIK
SLEGGDDDLPTCTLRSTTQCLTRPELYFMDVRSAINKGTDEGALTLITVRABID
LKVIGBEYQRNSIPIEKATTKDRGDYEMLVALLGEDDA"
FEATURES
source

```

```

ORIGIN
Alignment Scores:
Pred. No.:      1,77e-126      Length:      1159
Score:          1585.00      Matches:     312
Percent Similarity: 98.74%      Conservative: 1
Best Local Similarity: 98.42%      Mismatches:  4
Query Match:    98.39%      Indels:      0
DB:             8      Gaps:      0

US-10-690-564-2 (1-317) x ATU28415 (1-1159)

Qy      1 MetAlaThrLeuLysValSerAspSerValProAlaProSerAspAspAlaGlnGlnLeu 20
Db      6 ATGCGACCTTAAAGGTTCTGATTCGTTCCTGCTCTTGATGATGATGATGATGATGAT 65
Qy      21 ArgThrAlaPheGlnGlnLysTyrGlnArgArgAspLeuIleIleSerIleLeuAlaHis 40
Db      66 AGAACCGCTTTTGAAGAGATGGGTACGAAAGAGACTTGAATCATCATCTTGGCTCAC 125
Qy      41 ArgSerAlaGlnGlnArgLysValIleArgGlnAlaTyrHisGlnThrTyrGlnLysAsp 60
Db      126 AGAAGTGCTGAACAGAGAAAGTCATCAGGCAAGCATACCAAGAACTTACCGCAAGAAC 185
Qy      61 LeuLeuLysThrLeuAspLysGlnLysSerAspAspPheGlnArgAlaIleLeuLeuTyr 80
Db      186 CTTCTCAAGACTCTTGAACAAAGAGCTCTTGAAGATTTGAGAGAGACTATCTTGTTGTTGG 245
Qy      81 ThrLeuGlnProGlnArgArgAspAlaLeuLeuAlaAsnGlnAlaThrLysArgTyrThr 100
Db      246 ACTCTTGAAACCGGTAGAGCTGATGCTTTATTTGGTCAATGAAGCTCAAAAAGATGACT 305
Qy      101 SerSerAsnGlnValLeuMetGlnValAlaCysThrArgThrSerThrGlnLeuLeuHis 120
Db      306 TCAAGCAACCAAGTCTTATGGAAGTGTCTTGCAAGAGCATCAACGAGCTGCTTAC 365
Qy      121 AlaArgGlnAlaLysThrHisAlaArgTyrLysLysSerLeuGlnGlnAspValAlaHisHis 140
Db      366 GCTAGGCAAGCTTACCATCTCTGCTTCAAGAAAGTCTTGAAGAGAGCTTGCTCACAC 425
Qy      141 ThrThrGlyAspPheArgLysLeuLeuValSerLeuValThrSerTyrArgTyrGlnGly 160
Db      426 ACTACCGGTGACTTCGAAAGCTTTGGTTCTCTTGTTCATACCATACAGTACGAAAGA 485
Qy      161 AspGlnValAspMetThrLeuAlaLysGlnGlnAlaLysLeuValHisGlnLysIleLys 180
Db      486 GATGAAGTGAACATGACATTTGGCTAAGCAAGACTTGAAGCTGATGAGAAATCAAG 545
Qy      181 AspLysHisIleTyrAsnAspGlnAspValIleArgIleLeuSerThrArgSerLysAlaGln 200
Db      546 GACAAACACTTACATATATAGATGATGATTTAGATCTTGTCCACAAAGAGCAAAAGCTCAG 605
Qy      201 TLeaanaLathrPheanaaryTyrGlnaAspaPhisgljgluLuleuLysSerLeu 220
Db      606 ATCAATGCTACTTTTAAACCGTTTACCAAGATGATCATGGCGAGAAATTTCTCAAGACTCTT 665
Qy      221 GlnGlnGlnLysAspaAspaAspaAspaAspaAspaAspaAspaAspaAspaAspa 240
Db      666 GAGGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 725
Qy      241 ThrArgProGluLeuTyrPheValAspValLeuArgSerAlaIleAnlySerThrGlyThr 260
Db      726 ACAAGACCAAGACTTATCTTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 785
Qy      261 AspGlnGlnAlaLeuThrArgLeuValThrThrArgAlaGlnLysAspLeuLysValIle 280
Db      786 GATGAAGAGACACTCATTAATTTGATGACCAAGAGCTGAGATTGACTTGAAGCTCAT 845
Qy      281 GlnGlnGlnLysTyrGlnArgArgAsnSerIleProLeuGlnLysAlaIleThrLysAspThr 300
Db      846 GGAAGAGATACCAAGCGAGCAAGCATTCCTTTGAGAAAGCTATTACCAAGACACT 905
Qy      301 ArgGlyAspTyrGlnLysMetLeuValAlaLeuLeuGlnLysAspAspAla 317

```



Db 906 CGTGAAGATTACGAAGAAGATGCTCGTCGACCTTCTGGTGAATGATGCT 956

RESULT 8  
LOCUS ATANNEX 1206 bp mRNA linear PLN 13-JUL-1996  
DEFINITION A.thaliana mRNA for annexin.  
ACCESSION X99224  
VERSION X99224.1 GI:1429206  
KEYWORDS annexin; calcium binding protein; cytosolic protein.  
SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE 1  
Schantz, R., Schantz, M.L. and Houline, G.  
Unpublished  
JOURNAL 2 (bases 1 to 1206)  
AUTHORS Schantz, R.  
TITLE Direct Submission  
JOURNAL Submitted (09-JUL-1996) R. Schantz, Institut de Biologie  
Moléculaire des Plantes, CNRS, 12 Rue de General Zimmer, F- 67084  
Strasbourg, FRANCE  
Related sequences: Z47714 & Z33916.  
Location/Qualifiers  
1..1206  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/sub\_species="ecotype Columbia"  
/db\_xref="taxon:3702"  
/clone="FAFM80"  
/clone\_1b="1ambda Zap II"  
<1..947  
/function="calcium binding protein"  
/codon\_start=3  
/product="annexin"  
/protein\_id="CAA67608.1"  
/db\_xref="GI:1429207"  
/db\_xref="GCA:Q39001"  
/db\_xref="Uniprot/TREMBL:Q39001"  
/translation="TILKSDSVAPSDDAEQULTAFEGKGTNEDLIISILAHSAEQR  
KVIHQAVHETVYEDLTKTLDELKSLNDPEBALILMTLEPGSRDALINAEATKWTSSNQ  
VLMEEVACTRTSTQLHARQAVHARVYKSEEDVAHHTTDPFRKLVSIVTSRYEGDE  
VMNTLAKOEAKIVHEKIKDKNVDEDEVITLSTISKAQINATFNRRYODDHGEIILKSL  
EEGDDDDKFLALRSTIOCTTRPELFPVIVRSAINKGTDEBALTRIVTAAEIDLK  
VIGEEYQRNSIPLRAITTDKNCSDYGMVLVIFGBDA"

ORIGIN

Alignment Scores!

Pred. No.:	2, 03e-123	Length:	1206
Score:	1549.50	Matches:	307
Percent Similarity:	98.10%	Conservative:	2
Best Local Similarity:	97.46%	Mismatches:	5
Query Match:	96.18%	Indels:	1
DB:	8	Gaps:	1

US-10-690-564-2 (1-317) x ATANNEX (1-1206)

QY 3 ThtleuYvaIseRasPseValProAlaProSeRasPAsPaAgluGluLeuAArgTtr 22  
Db 3 ACTCTTAAGGTTCTGATCTGTTCCGCTCTTCGATGATCTGAGCAATTGAGAAC 62

QY 23 AlAPheGluGlyTrrpGlyThraaGluAAspleuIleIleSerIleleuIahIasRgSer 42  
Db 63 GCTTTGAAGAGATGGGCTACGAACGAGACTTGAATCAATCAATCTTGCTCAACAAGT 122

QY 43 AlAGlGluAArgTysValIleAArgAlaIaTyrHISgluThrTyrglyGluAAspleu 62  
Db 123 GCTGAACAGAGAAAGTCATCAGGCAAGCATACCAAGAACTTACGGCGAAGCCTTCTC 182

QY 63 LysThrleuAspIysGluLeuSerAsnAspPheGluAArgAlaIleleuLeuTrrHleu 82

Db 183 AAGACTTTGACAAGAGAGCTCTTAACGATTTGAGAGAGCATCTTGTGGACTCTT 242  
QY 83 GluProGlyGluAArgAspAlaIleuIleuAlaAaGluAlaThrTyraArgTrrPheSer 102  
Db 243 GAACCCGGGAGACGGATGCTTATTTGGTAAATGAGCTACAAAAGATGAGCTTCAAGC 302

QY 103 AaGluValIleuMetGluValAlaCySerThraArgTrrSerThrGluLeuIahIaArg 122  
Db 303 AACCAAGTCTTATGGAAGTTCCTTGCAACAGACATCAACGACCTGTCACCTAG 362

QY 123 GluAlaTyRHlaIaArgTyrlsYsSerIeugluGluAAspleuAlaIahIeThrThr 142  
Db 363 CAAGCTTACCAAGCTCGCTACGAAGAAGTCTTGAAGAGAGAGCTGTCACCACTACC 422

QY 143 GlYAAspPheAArgGlyLeuLeuValSerIeuvAlItrSerTyArgTrrGluGlyAspGlu 162  
Db 423 GGATGCTTACGAAGCTTTGGTTCTCTGTTACTCTATACAGGTACGAAGAATGA 482

QY 163 ValAsnMetThrleuAlaIysGluGluAlaIysLeuValIleGluIysIleYsAspLys 182  
Db 483 GTGAACATGACATTTGGCTTACGACAGAGTAAGCTGTCATGAGAAATCAAGACAAAG 542

QY 183 HsTyraAspAspGluAspValIleArgIleIeuserThraArgSerTyraAgluIleAsn 202  
Db 543 CACTCAATGATGAGAGATGTTATTAGAACTTGTCCACATTAAGCAAAAGCTCAGATCAAT 602

QY 203 AlAtrrPheAsnAArgTyrgluAspAspAspIleGlyGluIleleuYsSerIeuglu 222  
Db 603 GCTACTTTTAACCGTTACCAAGATATATCGCGAGAAATCTCAAGAGCTTACAGAA 662

QY 223 GlYAAspAspAspAspLysPheleuAlaIeuleuAArgSerThrIleGlySerleuThraArg 242  
Db 663 GGAAGATGATGATGACAGAGTCTCTTGCACCTTTGAAGTCAACCATTCAGTGTGACAAAG 722

QY 243 ProGluLeuTyrrPheValAspValIeuvArgSerAlaIleAsnTyThrGlyThraAspGlu 262  
Db 723 CCAGAGCTTACTTTGTCGATGCTTCTGCTTCCAGCAATCAAAAACCTGAACATGATGA 782

QY 263 GlYAAlaIeuvThraArgIleValThrThraArgAlaGluIleAspLeuValIleGlyGlu 282  
Db 783 GGAAGCATCTACTGAATTTGTGACCAACAGAGCTGATGACTTGAAGTCAATTGAGAG 842

QY 283 GluTyrgluAArgAAsnSerIleProIeugluGluValaIleThrTyraAspThraArgIy 302  
Db 843 GAGTACCGAGCGAGAGAACGATTCCTTGG---AAGATATTACCAAGAGCAATGTGGA 899

QY 303 AspTyrgluYsMetLeuValAlaIeuleuGlyGluAspAspAla 317  
Db 900 GATTACGGGAGAGATGCTTGTCCCAATTTTCGGGGAAGATGATGCT 944

RESULT 9  
LOCUS AY356355 954 bp mRNA linear PLN 01-SEP-2004  
DEFINITION Brassica juncea annexin mRNA, complete cds.  
ACCESSION AY356355  
VERSION AY356355.1 GI:38049639  
KEYWORDS

SOURCE  
ORGANISM  
Brassica juncea  
Brassica juncea  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Brassica.

REFERENCE 1 (bases 1 to 954)  
JAM, S.K. and KIRTI, P.B.  
Annexin cDNA from the floral bud of Brassica juncea  
Unpublished  
JOURNAL 2 (bases 1 to 954)  
AUTHORS Jami, S.K. and Kirti, P.B.  
TITLE Direct Submission  
JOURNAL Submitted (30-JUL-2003) Department of Plant Sciences, School of Life Sciences, University of Hyderabad, Gachibowli, Hyderabad, Andhra Pradesh 500046, India  
Location/Qualifiers

```

source
1. .954
/organism="Brassica juncea"
/mol_type="mRNA"
/db_xref="taxon:3707"
/issue_type="floral bud"
1. .954
/codon_start=1
/product="anexin"
/protein_id="AA10457.1"
/db_xref="GI:38049640"
/translation="MATLKVSSVSPSEDAEQLKSAFDGKGTNEELIISILAHRSAE
ORLIRQTHSEFEDLLKSELKELTSDPEALILMTLEBERDALVNEATKMTSS
NOVMEVACTRTSTQLHARQAYARFKSIIEEDVAHTTGDEFKILSVSSRYSG
EEVMTLAKQBAKLIHEKIKDKNYNDPFRILSTRKAOIINFTNNRYODNHEILK
SLBERGDDKPRGILRSTICTRPELXPYDVLSAINTKGTDEGALTRIVATRAEID
LKVIQGEYQRNRSIPLEKAIITKTRGDIEMKILALGEDDA"

ORIGIN

Alignment Scores:
Pred. No.: 6,61e-120 Length: 954
Score: 1507.00 Matches: 289
Percent Similarity: 98.11% Conservative: 22
Best Local Similarity: 91.17% Mismatches: 6
Query Match: 93.54% Indels: 0
DB: 8 Gaps: 0

US-10-690-564-2 (1-317) x AY356355 (1-954)

QY 1 MetAlaThrLeuIysValSerAspSerValProAlaProSerAspAlaGluGlnLeu 20
DB 1 ATGGCGACTCTTAAGGTTCTTCTTCTGTTCTCTCCCTCGAAGATGCTGACCAATTG 60
QY 21 ArgThrAlaheGluGlyTrpGlyThrAsnGluAspLeuIleIleSerIleLeuAlaHis 40
DB 61 AAAAGCCCATTTGGATGGATGGGATGACCAACGAGAAATGATCATATCAATCTGGCTCAC 120
QY 41 ArgSerAlaGluGlnArgLysValIleArgGlnAlaThrHisGluThrGlyGluAsp 60
DB 121 AGAAGTCTGACAGAGAGAGCTGATGCGCAACATACATGAAATCTTGGAGAGGAT 180
QY 61 LeuLeuIysThrLeuAspLysGluLeuSerAsnAspPheGluAlaIleLeuLeuTrp 80
DB 181 CTCTTAAGAGTCTTGAAGAGAACTTAACAAGCACTTCAGAGAGCATTGCTCTGG 240
QY 81 ThrLeuGluProGlyGluArgAspAlaLeuLeuAlaAsnGlnAlaThrLysArgTrpThr 100
DB 241 ACTCTTGAACCGGATGAACGTGATGCTTATTTGGTTATGATGAAGCTACCAAAAGATGACT 300
QY 101 SerSerAsnGlnValLeuMetGluValAlaCysThrArgThrSerThrGlnLeuLeuHis 120
DB 301 TCAGCAACCAAGTGTCTTAAGAGTGTCTTGCATGACCTCTACGACGCTTCTTCCAC 360
QY 121 AlaArgGlnAlaThrHisAlaArgTrpLysLysSerLeuGluGluAspValAlaHisHis 140
DB 361 GCTAGGCAAGCTTAACACGCTCGCTCAAGAGCTTATTAAGAGAGATGTCGCTCACAC 420
QY 141 ThrThrGlyAspPheArgLysLeuLeuValSerLeuValThrSerTrpArgTrpGlyGly 160
DB 421 ACCACCGGTACTTCAGAAAGCTTTGGTTCCTTGTGTGCTCATACAGTTCAGAGGG 480
QY 161 AspGluValAlaMetThrLeuAlaLysGlnGluAlaLysLeuValHisGluLysIleLys 180
DB 481 GAAAGAGTAAACATGACATTGGCAAAACAGAGGCTTAAGCTGATTCATGAGAAATCAG 540
QY 181 AspLysHisTrpAsnAspGluAspValIleArgIleLeuSerThrArgSerLysAlaGln 200
DB 541 GACAAGCATTAACAATGATGAAGATTTCATTAAGATTTCCTCAAGAGCAAAAGACAG 600
QY 201 ILAASnAlaThrPheAsnArgTrpGlnAspAspHisGlyGluGluIleLeuLysSerLeu 220
DB 601 ATCAATGCTACTTCAATCGCTTATCAAGATTAATCAACGCGCAGAAATCCCTCAAGAGCTT 660
QY 221 GluGluGlyAspAspAspAspLysPheLeuAlaLeuLeuArgSerThrIleGlnCysLeu 240

```

```

DB 661 GAGAAAGAGATGAAGAAGCAAGACTTCAGGCTGTGAGGCTCAACCATTCATGACTTG 720
QY 241 ThrArgProGluLeuTrpPheValAspValLeuArgSerAlaIleAsnLysThrGlyThr 260
DB 721 ACAAGACTGAGCTTACTTACTTGTGATGTTCTTCTGTTACAGCGATCAACAAACGGGACA 780
QY 261 AspGluGluValAlaLeuThrArgLysIleValThrThrArgAlaGluIleAspLeuValIle 280
DB 781 GACGAAGAGACTCTCACTGAATTTGGCCCAAGAGCTGAGATGACTTGAAGATCATT 840
QY 281 GlyGluGluTrpGlnArgArgAsnSerIleProLeuGluLysAlaIleThrLysAspThr 300
DB 841 GCAACAGAGTACCAAGAAAGAAACACATTCATTGAGAAACCATTAACCAAGACACT 900
QY 301 ArgGlyAspTrpGlyLysMetLeuValAlaLeuLeuGlyGluAspAla 317
DB 901 CGTGAAGATTACGAGAAAGATGCTCATGCGACTTCTCGTGAAGATGATGCT 951

RESULT 10
FL4D7 81513 bp DNA linear PLN 28-JUN-2000
LOCUS Sequence of BAC FL4D7 from Arabidopsis thaliana chromosome 1,
DEFINITION complete sequence.
ACCESSION AC021198
VERSION AC021198.2 GI:6957696
KEYWORDS HTG.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 81513)
Liu,S.X., Chan,A., Sakano,H., Yu,G., Lee,J.M., Lenz,C., Pham,P.,
Toriumi,M., Chin,C., Chlou,J., Choi,E., Chung,M., Gonzalez,A.,
Hwang,B., Liu,A., Vayberg,M., Alcafi,H., Brooks,S., Buehler,B.,
Chao,Q., Conn,L., Conway,A.B., Hansen,N.F., Johnson-Hopson,C.,
Khan,S., Kim,C., Lam,B., Miranda,M., Nguyen,M., Palm,C.J.,
Shim,P., Southwick,A., Davis,R.W., Becker,J.R., Federpiel,N.A.,
Theologis,A. and Yu,G.
The sequence of BAC FL4D7 from Arabidopsis thaliana chromosome 1
Unpublished
2 (bases 1 to 81513)
Theologis,A.
REFERENCE Direct Submission
AUTHORS Submitted (15-JAN-2000) Plant Gene Expression Center, 800 Buchanan
JOURNAL Street, Albany, CA 94710, USA
3 (bases 1 to 81513)
Theologis,A.
REFERENCE Direct Submission
AUTHORS Submitted (09-FEB-2000) Plant Gene Expression Center, 800 Buchanan
JOURNAL Street, Albany, CA 94710, USA
4 (bases 1 to 81513)
Theologis,A.
REFERENCE Direct Submission
AUTHORS Submitted (28-JUN-2000) Plant Gene Expression Center, 800 Buchanan
JOURNAL St., Albany, CA 94710, USA
COMMENT On Feb 11, 2000 this sequence version replaced gi:6693723.
The sequence is of BAC FL4D7 from Arabidopsis thaliana chromosome
1. The sequence does not represent the sequence of the entire
insert of this clone. It is shorter by 6954 bp because we submit
only the unique sequence of the clone. However, in order to
facilitate the joining of overlapping clones in the future for
creation of larger contigs, we provide small overlaps (200 bp)
between overlapping submitted clones. The 5' end of this sequence
overlaps by 200 bp to the 3' end of the sequence of the clone
FL504.

FEATURES
source
1. .81513
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/cultivar="Columbia"
/db_xref="taxon:3702"

```

gene /chromosome="1"  
/clone="F14D7"  
1553. .4999  
/gene="F14D7.1"  
join(1553. .4628,4713. .4999)  
/note="Contains similarity to receptor protein kinase-like protein from Arabidopsis thaliana gb|A161513. It contains a eukaryotic protein kinase domain pf|00069. EST gb|A1997574 comes from this gene."  
/codon\_start=1  
/evidence=not experimental  
/protein\_id="AAF79881.1"  
/db\_xref="GI:8778966"  
/translation="MGPAEKNIYDPRFLPIRITISCSISASATIAVANLLKWKSTP  
TNSKSLSWHDANTTSRSTSWGVSCSKRSISIELNITNTGICGTODPFILS  
NLAYVDLSMNLISGTTIPQGNLSKLIYFDLSTNHLGELISPSLGNLKNLTVLYLQ  
YLTYSVISELGNMESMTDLALSONKLTGSIIPSLGNLKNLMTVLYLYENLTGVIPEL  
GNMESMTDLALSONKLTGSIIPSLGNLKNLMTVLYLYENLTGVIPELGNMESMTLA  
LSQNKLTGSIIPSLGNLKNLTLTSLFONYLTGIPPLGNIESMIDLEISNNKLTGSI  
PSLGNLKNLTLTYLYENLTGVIPELGNMESMTDLALNNKLTGSISSPNNKLTGSI  
TYLYLYNTGVIPELGNMESMTDLALSONKLTGSIIPSLGNLKNLTVLYLYLQ  
SGAIPPEVANSHTLTLIDTNNPFGPEPTVCKGRKLQNTSIDVNHLESPILPKSLAD  
CKSLIRARFLGNKFTGDIPEAFGIYDPLNFDPSHNKFGELISSNWEKSPKGLIMS  
NNITGAIPELGNMESMTDLALSONKLTGSIIPSLGNLKNLMTVLYLYENLTGVIPEL  
GNMESMTDLALSONKLTGSIIPSLGNLKNLMTVLYLYENLTGVIPELGNMESMTLA  
LSQNKLTGSIIPSLGNLKNLTLTSLFONYLTGIPPLGNIESMIDLEISNNKLTGSI  
PSLGNLKNLTLTYLYENLTGVIPELGNMESMTDLALNNKLTGSISSPNNKLTGSI  
TYLYLYNTGVIPELGNMESMTDLALSONKLTGSIIPSLGNLKNLTVLYLYLQ  
ISICANTFYCTCRKQLQGNNTDPEIGENSIISVDGFKYQDIIESNEEDPVHL  
IGTGVSKYVRANLQDTIIAVKRLHDTIBEISKPVQKQFLNEVKALEIRHRNPHK  
LFGFGRHRTFLIYEMKESLNKLANDBEAKRLWTNRINNVAGVLAHEISYMHAD  
RITPIYHRTDSSGNLIDNDYTAKISDGFPAKLKTDSSNWSAVAGTYVAPEFAYT  
MKYTEKDYVSPGVILIELIKGHPGLVSSLSSPGBALSLRSIDERYLEBRQGNR  
EKLKNVEMALLCLQANPEBSPMTLSITTFSS"  
5917. .7552  
/gene="F14D7.2"  
join(5917. .5992,6589. .6734,6821. .7552)  
/gene="F14D7.2"  
/note="Identical to annexin (Annx1) mRNA from Arabidopsis thaliana gb|AF083933. It contains an annexin domain  
pf|00191. ESTs gb|H76460, gb|Z18518, gb|Z26190,  
gb|N96455, gb|Z47714, gb|T41940, gb|T43657, gb|N59995,  
gb|R30014, gb|T22046, gb|H37398, gb|H77008, gb|R29768,  
gb|H36260, gb|Z17514, gb|W43175, gb|T76739, gb|AA712753,  
gb|H76134, gb|T42209, gb|H36536, gb|A1988553, gb|Z32565,  
gb|AA597533, gb|A1100145 and gb|A110054 come from this gene."  
/codon\_start=1  
/evidence=not experimental  
/protein\_id="AAF79882.1"  
/db\_xref="GI:8778967"  
/translation="MATIKYSDSPAPSDAEDQRTAFEGQCTNEDLIISILHRSAB  
QKRVTRQAHETYGEDILKTLDKELSDNERALLTLTFEGEDALLANLAEATKWTSS  
NOVLMEVACTRTSTQLHARQAVHARYKSLDEVAHHTTGDRKLVLVLSVREG  
DEVNMTLAKQEAFLVHEKIKDKHYNDEVDVRIISTRSQINATFNRYQDDBEELIK  
SLEGGDDDDFALRLSTIOCLTRPELYFDVARSAINKGTGDEGALTRIVTFRABID  
LKVIIGEYORRNSIPLKATTKDTRGDYEMVALLGEBDA"  
complement(7937. .10730)  
/gene="F14D7.3"  
complement(join(7937. .8041,8115. .8364,8443. .8606,  
8757. .8919,9397. .10730))  
/gene="F14D7.3"  
/note="Contains similarity to a putative protein CAB78009  
gi|7267527 from Arabidopsis thaliana BAC T32A17  
gb|A161513. It contains Pumilio-family RNA binding  
domains pf|00806."  
/codon\_start=1  
/evidence=not experimental  
/protein\_id="AAF79883.1"  
/db\_xref="GI:8778968"  
/translation="MKKNQKKRQKSVNSISAGCKKQCLRLCYINGEKTKYKNT  
TSHLRVTHATFTSLISQTSICYITPKLLSPCLGFTLSSISFVNCINYSVM  
FGRFPGFGFNGDFNRGASDHRPLGSDGFLPDLTNPFLKNHKSVEALDLCKKLHK

misc\_feature  
gene  
CDS  
/evidence=not experimental  
/protein\_id="AAF79884.1"  
/db\_xref="GI:8778969"  
/translation="MAKATVAIILCFTLIGEEIFSGVFAECCRGRFKULLFSVNDYL  
GSSPHTSINIANLFCBGLIGCSTSTSPFLSLKKHAKFLTEKVDADDEKXKND  
EKXGDDPYVR"  
complement(34011. .36183)  
/gene="F14D7.5"  
complement(join(34011. .34115,34201. .34450,34532. .34788,  
34873. .35035,35279. .36183))  
/gene="F14D7.5"  
/note="Containing similarity to a putative RNA-binding protein AAR18550 gi|6587864 from Arabidopsis thaliana BAC T1111 gb|AC012680. It contains Pumilio-family RNA binding domains pf|00806."  
/codon\_start=1  
/evidence=not experimental  
/protein\_id="AAF79885.1"  
/db\_xref="GI:8778970"  
/translation="WEIFNFGASGSHRRLPDFGSGFQSLDTNPFLLKNQYRNSVEA  
LELCKRLNMGISCDMSIWTKEBPRVDGFGAKTLHESGFQNLTGASQIHDF  
RNFSSVRVONNNHFGVSPGFLIGDFNPGFEMMAFKHDFLIDHINEPIKRS  
PLRGNDVAKSGIMEGIRVSOIILAMESGASYPDEPIKNGDLIDLVNVEIIGSV  
NIMARDQIGRALQLVGEGTLDKSVFLIEDVHVEMLPNDLYIYQKLVVSD  
BGRMTIVSLTSEKPRELTKICLNTNGTGVQWITVYTKQOIALYKALREGFLVY  
NDSNGHVAQSCLEFLVPNDKRLRLTKSHIISFLFLITVSSPECCETKQFVYA  
ATEYCAQLATHQGYCYVLCQSLINTVGLQHEKLVAEISRDSLRISQDPGNVVOCLI  
DOVSSVNLLEPFRTHCIELATQKFSHVIEKCLARKYESRAEIVRELSYFNFOEL  
QDPYANVYIQTLASTVKAVARLVEKYRFQKLSNPYCKKIFSKTILKK"  
44612. .48698  
/gene="F14D7.6"  
join(44612. .44780,45014. .45403,45501. .45644,45676. .46134,  
46297. .46476,46633. .47029,47113. .47505,47737. .48082,  
48145. .48411,48465. .48541,48632. .48698)  
/gene="F14D7.6"  
/note="Strong similarity to a mutator-like transposase  
gi|4063759 from Arabidopsis thaliana BAC T1444  
gb|AC005561."  
/codon\_start=1  
/evidence=not experimental  
/protein\_id="AAF79886.1"  
/db\_xref="GI:8778971"  
/translation="MSDIYGVARVLYTLVYLLAVYSILCTICEKPRLTCTVSDKINA  
KAKIPENSRLKGVRLREDGDFAKYVMAVKEKLAGEDDIELYQWPPMMGPPDRR  
ANPIDLIDEDWTLFMAIRADIDEVHPVKIIRGNSRVNSVRSNLDVGLTLEIS  
DKYNSAETRAVAGTALTRMLRNISAHLSPIYLYATVYLYQVYMDQWERTTA  
MYRAMNTETLMGGAGTGVPLDTPEPDSIPISKEKEKYDDDEREDRLFEMWRT  
QOQEBEMCATOMTIGATOIDATNVGDTGAALANGSGDIFKNGRDELOLTLA  
RAEIOBCTQVAYALSSSDSTRMPSTISVSFSFSDLDLMGRTSSNSGCTIDGCK  
IMTIDERTTKEIGASTNTRVSEVDNEMAVDITVPVLYLRDAPPTFDDEKEDAL  
HNAKYADNEGDIYIGRIFKSKEDCATKLAHAHMHFPIYASGPAITVAACVSD  
TGPWRYATKLPDSDMFEVRSYTYQHTCSVDARQDFHQAQSTVVIQKLMRTKYIGVR



```
Db 274 GCTTTCATCTCTGCTGAACGTGATGCCCTTTGGCTATGACCAACCAAGGTGAGCT 333
Qy 101 SerSerAnlnlValleuMetGluValAlaCysThrArgThrSerThrGlnLeuHis 120
Db 334 TCAGCAATAGAGGTCTTATGGAATAAGCTTGACAGAGGTCTCCAAACCAACGTCTTAC 393
Qy 121 AlArgGlnAlaTyrHisAlaArgTyrLysSerLeuGlnGluAspValAlaHisHis 140
Db 394 GCAGGAGGCTTATCATCTGCTTATTAAGAAGTCGCTTAAGAAGATGTTGCTCATCAC 453
Qy 141 ThrThrGlyAspPheArgLysLeuLeuValSerLeuValThrSerTyrArgTyrGlnGly 160
Db 454 ACCACTGGCGCACTTCCCTAGCTCCCTCTACTCTACTCTAGTGAAGTCAATACATATGAGGGA 513
Qy 161 AspGlnValAsnMetThrLeuAlaLysGlnGluAlaLysLeuValHisGlnLysHis 180
Db 514 GAGGAGGTGAACATGAATCTGGCGAAAACAGAGCGCAAGTTCCTCATGAGAAATTTCA 573
Qy 181 AspLysHisTyrAsnAspGluAspValIleArgIleLeuSerThrArgSerLysAlaGln 200
Db 574 GACAAAGCTTACAGATGATGACGATGCTATAGGGTTTGGCTACAGAGCAAGGACAG 633
Qy 201 IleAsnAlaThrPheAsnArgTyrGlnAspAspHisGlyGlnGlnLysSerLeu 220
Db 634 ATCAATGCAACTCTGAAATCACTACAAAATGAATATGAAATGACATTAACCAAGACTTG 693
Qy 221 GlnGlnGluLysAspAspAspLysPheLeuAlaLeuLeuArgSerThrIleGlnCysLeu 240
Db 694 --AAGCTATCTTAAGATGAGTCTTGCACTCTAAGTCCACAGAGAGTCTTG 750
Qy 241 ThrArgProGluLeuTyrPheValAspValLeuArgSerAlaIleAsnLysThrGlyThr 260
Db 751 GTCATTCGCGAAAGATATTTTGAGAGGTTCTTCGCTAGCATCATATAGACAGGAGAC 810
Qy 261 AspGlnGluValaLeuThrArgIleValThrThrArgAlaGlnLysLeuValIle 280
Db 811 GATGAAGAGAGCTCTTACTAGAGTGTTCGACATGAGGCTGAGGTTGATCTTAAGATCAT 870
Qy 281 GlnGlnGluLysThrArgArgAsnSerIleProLeuGlnLysAlaIleThrLysAspThr 300
Db 871 GCAGATAGTACCAAGCAAGAAACAGTCCACTGCTGCCCATGTGCAAGGACACT 930
Qy 301 ArgGlyAspTyrGlnLysMetLeuValAlaLeuLeuGly 313
Db 931 CATGAGACTATGAAAAATTGCTGCTGACTTCTTGACAGA 969

RESULT 12
GHU73746 948 bp mRNA linear PLN 24-JAN-2003
LOCUS Gossypium hirsutum annexin (Annh1) mRNA, partial cds.
DEFINITION U73746
ACCESSION U73746.2 GI:15214409
VERSION 1
KEYWORDS
SOURCE Gossypium hirsutum (upland cotton)
ORGANISM Gossypium hirsutum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
REFERENCE 1 (bases 1 to 948)
AUTHORS Delmer,D.P. and Potlikha,T.S.
TITLES Structures and functions of annexins in plants
JOURNAL Cell. Mol. Life Sci. 53 (6), 546-553 (1997)
MEDLINE 97374485
PUBMED 9230934
REFERENCE 2 (bases 1 to 948)
AUTHORS Potlikha,T.S. and Delmer,D.P.
TITLES Direct Submission
JOURNAL Submitted (09-OCT-1996) Plant Sciences, Hebrew University, Givat
Ram Campus, Jerusalem 91904, Israel
COMMENT On Aug 20, 2001 this sequence version replaced gi:1843524.
FEATURES
source 1..948
```

```
/organism="Gossypium hirsutum"
/mol_type="mRNA"
/strain="Acala Sj-2"
/db_xref="taxon:3635"
/tissue_type="cotton fibers"
/dev_stage="21 days post-anthesis"
<1..948
/gene="Annh1"
<1..948
CDS
/gene="Annh1"
/function="evidence suggests that annexins may bind to and
influence activity of callose synthase"
/notes="contains 4 repeats characteristic of other
annexins; shows evidence of conserved calcium-binding
domains; by comparison with other plant annexins as well
as with the size of native cotton fiber annexin, this
clone probably lacks 5' coding sequence for only the
N-terminal Met; calcium binding protein"
/codon_start=1
/product="annexin"
/protein_id="AAB67993.2"
/db_xref="GI:15214410"
/translation="ATLTVPPTVPSVSDCEQLRFAEGWGTEGLIIDLIGHRNAEQ
KMLIRKYAETGEBDLKALDKELSDNERLVLALMDPAERDALANAEATKRTSSN
OVLAMEIACRSANQOLHARQAVHARVKSLEEDVAHHTGDFHKLLELVSSRYEGE
EVMNLATKEALKEKTSNKAISDDVIRVATRSKQINATLHNYKREYGNIDNKO
LKADKDEFLALRSTVRCIVPEYKFKVLRIALINRGTEBGLTRVVCRAEVDLK
VTADRYORNSVPLTRAIKDTKHGEKLLVLAQVHN"

ALIGNMENT Scores:
Pred. No.: 4.5e-93 Length: 948
Score: 1193.50 Matches: 229
Percent Similarity: 85.26% Conservative: 37
Best Local Similarity: 73.40% Mismatches: 45
Query Match: 74.08% Indels: 1
DB: 8 Gaps: 1

US-10-690-564-2 (1-317) x GHU73746 (1-948)
Qy 2 AlatrPheLysValSerAspSerValProAlaProSerAspAspAlaGlnLeuArg 21
Db 1 GCCACTCTTACAGTCCGCCACGACAGTCTTCAAGTCTGAAAGATGTGCAACACTAAGA 60
Qy 22 ThrAlaPheGlnGlyThrGlyThrAsnGluAspLeuIleSerIleLeuAlaHisArg 41
Db 61 AAAGCTTTTTCAGAGATGGGGAACATAAGAGGCTTAATCATGATATATGAGTCAACA 120
Qy 42 SerAlaGlnGluArgLysValIleArgGlnAlaTyrHisGlnThrTyrGlyGluAspLeu 61
Db 121 AATGGGAGCAACGAACCTGATTCGAATAAACCCTACGCGAAGAACTATGAGAGATCTC 180
Qy 62 LeuLysThrLeuAspLysGlnLeuSerAsnAspPheGlnArgAlaIleLeuLeuThr 81
Db 181 CTCAAAGGCACTAGACAAGAACTCTGCAATGCTTGAAGAGGCTGTTCTGCTTGAGCT 240
Qy 82 LeuGlnProGlyGluLysAspAlaLeuLeuAlaAsnGluAlaThrGlyAspThrPheSer 101
Db 241 TTTGATCTCTGCAAGCTGATGCCCTTTTGGCTATGAAAGCCACCAAAAGGTGACTTCA 300
Qy 102 SerAsnGlnValleuMetGluValAlaCysThrArgThrSerThrGlnLeuHisAla 121
Db 301 AGCATCAAGTCTTATGGAATATACCTGCGCAAGGTCGCCAACCAACTGCTTACGGA 360
Qy 122 ArgGlnAlaTyrHisAlaArgTyrLysLysSerLeuGlnGluAspValAlaHisHisThr 141
Db 361 AGCGAGGCTTATCATGCTCGTTATAAGAAAGTGGCTTGAAGAGAGAGTCTCATACAG 420
Qy 142 ThrGlyAspPheArgLysLeuLeuValSerLeuValThrSerTyrArgTyrGlnGly 161
Db 421 ACTGGGAACTTCATTAAGCTCTCTACTCTAGAGATGCTATACAGATATAGGAGAGG 480
Qy 162 GluValAsnMetThrLeuAlaLysGlnGluAlaLysLeuValHisGlnLysHisAsp 181
```



LOCUS AF006197 1112 bp mRNA linear PLN 22-AUG-2001  
 DEFINITION Lavatera thuringiaca amexin (AmxL1) mRNA, complete cds.  
 ACCESSION AF006197  
 VERSION AF006197.1 GI:2459925  
 KEYWORDS  
 SOURCE Lavatera thuringiaca  
 ORGANISM Lavatera thuringiaca  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eucosids II; Malvales; Malvaceae; Malvoideae; Lavatera.  
 1 (bases 1 to 1112)  
 Vazquez-Tello, A. and Uozumi, T.  
 Cloning and characterization of a Lavatera thuringiaca cDNA  
 encoding an amexin whose expression is stimulated by low  
 temperature  
 Plant Physiol. (1997) In press  
 2 (bases 1 to 1112)  
 Breton, G., Vazquez-Tello, A., Danyluk, T. and Sarhan, F.  
 Two novel intrinsic amexins accumulate in wheat membranes in  
 response to low temperature  
 Plant Cell Physiol. 41 (2), 177-184 (2000)  
 20255875  
 MEDLINE 10795312  
 PUBMED 3 (bases 1 to 1112)  
 REFERENCE Vazquez-Tello, A. and Uozumi, T.  
 AUTHORS Submitted (31-MAY-1997) Biological Sciences, Universite du Quebec,  
 TITLE Direct Submission  
 JOURNAL Montreal, C.P. 8888, Succ. Centre-Ville, Montreal, Quebec H3C 3P8,  
 Canada

# FEATURES

source Location/Qualifiers  
 1. 1112  
 /organism="Lavatera thuringiaca"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:61660"  
 /tissue\_type="leaves without petioles"  
 /note="From England"  
 1. 1112  
 /gene="AmxL1"  
 2. 952  
 /function="binds acidic phospholipids in a  
 calcium-dependent manner"  
 /note="calcium-dependent, phospholipid binding protein;  
 contains four characteristic repeats found in both animal  
 and plant amexins"  
 /codon\_start=1  
 /product="amexin"  
 /protein\_id="AA571830.1"  
 /db\_xref="GI:2459926"  
 /translation="MATLTVPSTLPVSSEDCQIRKAFSGWGTNEDLIINITLGHNRAD  
 ERNSIRKAYETHEGDELKALDELSENPERLVILWLPDPERDALLANATKRTSS  
 NOVIMEIACRSSSDOLLRARQAVHVRKKSLSEDAVHHTGGDRKLLPLVSSRYEG  
 DEYNMTLATREATLIEKISNKRYSDDDPVRVLTATRSKQINRNLHYHNGEYATDINK  
 DLKADPRDEFLALRSTVKLVPEKRYFEVRLALINKGIDEGALTRVAVSTRAEVDL  
 KIIADBYQRNSVPLTRAIKVDINQDEKLLVLAGEVA"

## ORIGIN

### Alignment Scores:

Pred. No.: 2.83e-91 Length: 1112  
 Score: 1173.50 Matches: 225  
 Percent Similarity: 84.54% Conservative: 43  
 Best Local Similarity: 70.98% Mismatches: 48  
 Query Match: 72.84% Indels: 1  
 DB: 8 Gaps: 1

US-10-690-564-2 (1-317) x AF006197 (1-1112)

QY 1 MetAlAThrLeuIySValSerAspSerValProAlAProSerAspAspAlaGluGlnLeu 20  
 Db 2 ATGGCTACTCTTACAGTCCCTCCACACCTCCGTCAGTGTCTGAAGTTCTGAAACAAC 61  
 QY 21 ArgThrAlaPheGluGlyTTPGlyThrAsnGluAspLeuIleIleSerIleLeuAlaHis 40

Db 62 AGGAAAGCCTTCTCAGATGAGGGAAGCTAATGAGAGCTTAATCATTAATATTGGGTCA 121  
 QY 41 ArgSerAlaGluGlnArgIySValIleArgGlnAlaIyRnHLeuGluThrTyGluGluAsp 60  
 Db 122 CGAATGCGGACGACGAAACGAACTCGATTCGAAAAGCTTATCTGAACCATGAGAAAGAT 181  
 QY 61 LeuLeuIySThrLeuAspIySgluSerAsnAspPheGluArgAlaIleLeuLeuTrp 80  
 Db 182 CTCCTCAAGCAGCTGGACAGGAAGACTCTCAATGACTTTGAGAGGCTGTGCTTGG 241  
 QY 81 ThrLeuGluProGlyIyGluArgAspAlaLeuLeuAlaAsnGluAlaThrIyAspTrp 100  
 Db 242 ACTCTTGATCCCTCGAAGGTATGACACTTTTGGGAAAGAACCCAAAGGGGAGCT 301  
 QY 101 SerSerAsnGluValLeuMetGluValAlaCysThrArgThrSerThrGlnLeuLeuHis 120  
 Db 302 TCAGCAATCAGCTAATTATGAAATATGAGCTTCAGAAAGTCTTTCGACCACCTGCTGCC 361  
 QY 121 AlaArgGlnAlaTyRnHLeuArgTyRnIySLeuSerLeuGluGluAspValAlaHisHis 140  
 Db 362 GCGAGGCAAGCTTATCATGATGCTGTTATGAAGAAATCGCTTGAAAGAGATGTTGCCATC 421  
 QY 141 ThrThrGlyAspPheArgIySLeuLeuValSerLeuValThrSerTyRnArgTyRnGlu 160  
 Db 422 ACAACTGGGAGCTTCGTAAGCTTCCTCACTTGATGATTCATACAGATACAGAGGA 481  
 QY 161 AspGluValAsnMetThrLeuAlaIySglnGluAlaIySLeuValHisGluIySLeuIyS 180  
 Db 482 GATGAAGTAAACATGACTGTGGCGAAACAGAGGCCAACTTACTCAGTAGAAATCTCA 541  
 QY 181 AspIySHLeuTyRnAsnAspIySAspValIleArgIleLeuSerThrArgSerIySAlaGln 200  
 Db 542 AACAAAGCTTACAGATGATGACATGTCATCAGGGTTTGGCTACGAAAGCAAGTACAG 601  
 QY 201 IleAsnAlaThrPheAsnArgTyRnIySAspAspHisGlyGluGluIleLeuIySLeu 220  
 Db 602 ATCAACGAAACGCTTATATCATCAACAAATGAATACGCAACTGATATATAACAGGACCTG 661  
 QY 221 GluGluGluIySAspAspAspIySAspPheLeuAlaLeuLeuArgSerThrIleGlnCysLeu 240  
 Db 662 ---AAGGCTGACCCCTAAGATGATGCTTCCTGCACTGCTAAGGTCCACAGTGAAGTCTTG 718  
 QY 241 ThrArgProGluLeuLeuTyRnIySAspValLeuAspSerAlaIleAsnIySThrGlyThr 260  
 Db 719 GCTACCCCTGAAAGATATTGAGAAAGGTTCTTCGCTCAATCAATTAACAGAGGACG 778  
 QY 261 AspGluGluValAlaLeuThrArgIleValThrThrArgAlaGluIleAspLeuIySValIle 280  
 Db 779 GATGAAGAGGCTCTTACGAGGGTGTTCACACAGGGCTGAGGGTGTGATCTTAAGATCA 838  
 QY 281 GlyGluGluTyRnIySAspArgAsnSerIleProLeuGluIySValAlaIleThrIySAspTrp 300  
 Db 839 GCAGATGATACACAGGAAAGAACAGTGTCCCACTGATCTCGTCAATTTCTCAAGAGACT 898  
 QY 301 ArgGlyAspTyRnGluIySLeuLeuValAlaLeuLeuGluIySAspAspAla 317  
 Db 899 AATGAGACTACGAAAAATTTGCTGCTGTGTACTTCTGAGAGAGTGAAGGCT 949

RESULT 15  
 BD236034 1293 bp DNA linear PART 17-UTL-2003  
 LOCUS BD236034  
 DEFINITION Materials and method for modification of plant cell wall  
 POLYSACCHARIDES.  
 ACCESSION BD236034  
 VERSION BD236034.1 GI:33045804  
 KEYWORDS UP 2002527056-A/56.  
 SOURCE Pinus radiata (Monterey pine)  
 ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.  
 1 (bases 1 to 1293)  
 Bloksberg, L.N.  
 TITLE Materials and method for modification of plant cell wall





**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 22, 2005, 06:27:11 ; Search time 665 Seconds

(without alignments)  
2821.892 Million cell updates/sec

Title: US-10-690-564-2

Perfect score: 1611

Sequence: 1 MATLKVSDSPAPSDAEQL.....KOTRGDYEXKLVALLGSDDA 317

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 4390206 segs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=framed\_p2n.model -DEV=x1h  
-Q=/cgml\_1/USPTO.spool/US10690564/runat\_19082005\_172306\_25282/app\_query.fasta\_1.455  
-DB=N\_Geneseq\_16Dec04 -OPMT=fastcap -SUFFIX=ring -MINMATCH=0.1 -LOOPEXT=0  
-LOOPEXT=0 -UNITS=bites -START=1 -END=-1 -MATRIX=blonsum62 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USR=US10690564.CCGN\_1\_1\_470.0runat\_19082005\_172306\_25282 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -MARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N\_Geneseq\_16Dec04: \*  
1: geneseqn1980s: \*  
2: geneseqn1990s: \*  
3: geneseqn2000s: \*  
4: geneseqn2001as: \*  
5: geneseqn2001bs: \*  
6: geneseqn2002as: \*  
7: geneseqn2002bs: \*  
8: geneseqn2003as: \*  
9: geneseqn2003bs: \*  
10: geneseqn2003cs: \*  
11: geneseqn2003ds: \*  
12: geneseqn2004as: \*  
13: geneseqn2004bs: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1611	100.0	954	12	ADJ63871 Plant lip
2	1611	100.0	954	12	ADN74566
3	1611	100.0	1192	3	AACN34418 Arabidops
4	1611	100.0	1212	3	AAC34549 Arabidops
5	1566	97.2	1182	3	AAC33439 Arabidops

6	1202.5	74.6	1184	13	ADJ63461 Cotton cd
7	1185.5	73.6	2588	3	AAJ67100 Eucalyptu
8	1171.5	72.7	1293	3	AAJ67128 Pinus rad
9	1171.5	72.7	1293	4	AAJ44756 Annexin-1
10	1171.5	72.7	1293	10	ADJ94943 Programme
11	1132.5	70.3	1058	10	ADJ94854 Programme
12	1106.5	68.7	1157	10	ADJ73996
13	1067	66.2	954	12	ADN74790
14	1067	66.2	1156	3	AAJ9728
15	1067	66.2	1160	3	AAJ40038
16	1045	64.2	951	6	ABJ66078 Arabidops
17	1034.5	64.2	951	3	AAJ41736
18	988	61.3	1025	3	AAJ49738
19	960	59.6	1027	3	AAJ37764 Arabidops
20	866.5	53.8	789	3	AAJ67131 Pinus rad
21	866.5	53.8	789	3	AAJ67098 Eucalyptu
22	866.5	53.8	789	4	AAJ44759 Annexin-1
23	866.5	53.8	789	10	ADJ94946 Programme
24	793	49.2	589	13	ACN58684
25	785	48.7	602	13	ACN57334
26	778.5	48.3	954	12	ADJ07773
27	775	48.1	591	13	ACN60265
28	760	47.2	600	13	ACN60840
29	757.5	47.0	2132	3	AAJ67099
30	742	46.1	618	13	ADJ63464
31	738	45.8	586	13	ACN60745
32	660	41.0	627	3	AAJ67101 Eucalyptu
33	659	40.9	608	13	ACN53624
34	639	39.7	570	13	ACN58229
35	624.5	38.8	420	3	AAJ37667 Arabidops
36	602.5	37.4	966	6	ABJ44750 Arabidops
37	602.5	37.4	1159	3	AAJ45661 Arabidops
38	602.5	37.4	1160	3	AAJ42551 Arabidops
39	573	35.6	704	4	AAJ67132 Pinus rad
40	573	35.6	704	4	AAJ44760 Annexin-1
41	573	35.6	704	10	ADJ94947 Programme
42	563	34.9	704	3	AAJ67096
43	520.5	32.3	484	3	AAJ67129 Pinus rad
44	520.5	32.3	484	4	AAJ44757 Annexin-1
45	520.5	32.3	484	10	ADJ94944 Programme

#### ALIGNMENTS

RESULT 1	
ID	ADJ63871 standard; DNA; 954 BP.
XX	
AC	ADJ63871:
XX	
DT	06-MAY-2004 (first entry)
XX	
DE	Plant lipid metabolism protein OO-10 gene SEQ ID NO:73.
XX	
KW	de; gene; plant; lipid metabolism protein; LMP; seed storage compound;
XX	
XX	transgenic plant.
XX	
OS	Unidentified.
XX	
FH	Key
FT	CDS
FT	Location/Qualifiers
FT	/tag= 1.954
FT	/product= "OO-10"
XX	
XX	WO2004013304-A2.
XX	
PD	12-FEB-2004.
XX	
XX	04-AUG-2003; 2003WO-US024364.
PF	
XX	02-AUG-2002; 2002US-0400803P.
PR	
XX	
PA	(BADI ) BASF PLANT SCI GMBH.

XX Miltendorf V, Haertel HA, Bauer J, Oswald O;  
PI  
XX  
DR WPI; 2004-157121/15.  
DR P-PSDB; ADJ63872.  
XX  
PT New lipid metabolism proteins and nucleic acids, useful in producing  
transgenic plants with increased levels of seed storage compound, e.g.  
PT lipid, a fatty acid, a starch or a seed storage protein.  
XX  
PS Claim 1; SEQ ID NO 73; 115pp; English.  
XX  
CC The invention relates to novel isolated lipid metabolism proteins (LMP)  
CC and encoding nucleic acids comprising a polynucleotide sequence encoding  
CC a polypeptide that functions as a modulator of seed storage compounds in  
CC a plant. The LMP nucleic acid is useful in producing transgenic plants  
CC with increased levels of seed storage compound, e.g. lipid, a fatty acid,  
CC a starch or a seed storage protein, as markers for specific regions of  
CC the genome and for evolutionary and protein structural studies. The  
CC present sequence represents an LMP nucleic acid of the invention.  
XX  
SQ Sequence 954 BP; 288 A; 198 C; 229 G; 239 T; 0 U; 0 Other;  
XX  
Alignment Scores:  
Pred. No.: 2,06e-153 Length: 954  
Score: 1611.00 Matches: 317  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 12 Gaps: 0  
US-10-690-564-2 (1-317) x ADJ63871 (1-954)  
QY 1 MetAlaThrIleuLysValSerAspSerValProAlaProSerAspAlaGluGlnLeu 20  
DB 1 ATGGCAGCTCTTAAGGTTCTGATTCCTGCTCCGCTCTTATGATGCTGCAATTG 60  
QY 21 ArgThrAlaPheGluGlyTrpGlyThrAsnGluAspLeuIleSerIleLeuAlaHis 40  
DB 61 AGAACCGCTTTTAAGGATGGGTACGACGAGACTTGTATCATCATCTTGGCTCAC 120  
QY 41 ArgSerAlaGluGlnArgLysValIleArgIleAlaThrHisGluThrTyrglyLysAsp 60  
DB 121 AGAAGTGCTGAACAGAGAAAGTCATCAGCAGCAATACCAAGAACTCAGCGCAAGAC 180  
QY 61 LeuLeuLysThrIleuLysAspLysGluLeuSerAsnAspPheGluArgAlaIleLeuLeuTrp 80  
DB 181 CTTCTCAAGACTCTTGAACAAGAGCTCTTAACGATTTTCAGAGACTATCTTGTGTGG 240  
QY 81 ThrLeuGluProGlyLysArgAspAlaLeuLeuAlaAsnGluAlaThrLysAspTrpThr 100  
DB 241 ACTCTTGAAACCCGGTAGCGTGAATGCTTTATTTGGCTTAATGAACTTCAAAAAGATGACT 300  
QY 101 SerSerAsnGluValIleuMetGluValAlaCysThrArgThrSerThrGlnLeuLeuHis 120  
DB 301 TCAAGCAACCAAGCTTTATGAAAGTTGCTTGCAACAAGCAATCAACGAGCTGCTTCAAC 360  
QY 121 AlaArgGlnAlaTyrrHisAlaArgTyrrLysSerLeuGluLysAspValAlaHisHis 140  
DB 361 GCTAAGCAACCTTAACATGCTCTCTCAAGAAAGCTCTTGAAGAAGACGTTGCTCACAC 420  
QY 141 ThrThrGlyAspPheArgLysLeuLeuValSerLeuValThrSerTyrrArgTyrglyLys 160  
DB 421 ACTACCGGTACTCTTCAAGAAAGCTTTGGTTCTCTTGTTACCTCATACAGTACGAAGAA 480  
QY 161 AspGluValAlaMetThrLeuAlaLysGlnGluAlaLysLeuValHisGluLysIleLys 180  
DB 481 GATGAAGTGAACATGACATGGCTTAAGCAAGAGCTTAAGCTGATGAGAAATCAAG 540  
QY 181 AspLysHisTyrrAsnAspGluAspValIleArgIleLeuSerThrArgSerLysAlaGln 200  
DB 541 GACAAACACTAACAATGATGATGAGATGTTATTAGATCTTTGCCAAGAAAGCAAGCTCAG 600

QY 201 IleAsnAlaThrPheAsnArgTyrglnAspAspHisGlyGluGluIleuLysSerLeu 220  
DB 601 ATCAATGCTACTTTTAAACGTTTACCAAGATGATCATGGCGAGAAATCTCAAGAGCTCT 660  
QY 221 GluGluGlyAspAspAspAspLysPheLeuAlaLeuLysSerThrIleGlnCysLeu 240  
DB 661 GAGGAAGGAGATGATGATGACAAAGTTCTTGCACTTTTGAGAGTCAACCATTCAGTCTTG 720  
QY 241 ThrArgProGluLeuLysThrPheValAspValLeuAspSerAlaIleAsnLysThrGlyThr 260  
DB 721 ACAAGACCAAGCTTTACTTGTCCATGTTCTTCATGCAATCAACAAACTGGAACT 780  
QY 261 AspGluGlyAlaLeuThrArgIleValThrThrArgAlaGluIleAspLeuValIle 280  
DB 781 GATGAAGAGCACTCATGATTTGTGACCAAGAGCTGAGATGATTCAGAGTCAAT 840  
QY 281 GlyGluGlyTyrglnArgArgAsnSerIleProLeuGluLysAlaIleThrLysAspThr 300  
DB 841 GGAGAGGAGTACCAAGCGCAGAAAGCAATTCCTTGGAGAAAGCTATTAACAAAGACACT 900  
QY 301 ArgGlyAspTyrglyLysMetLeuValAlaLeuLeuGlyLysAspAspAla 317  
DB 901 CGTGAAGATTACGAGAAGATGCTCGTCCGACTTTCGGTGAAGATGATGCT 951  
RESULT 2  
ADN74566  
ID ADN74566 standard; cDNA; 954 BP.  
XX  
AC ADN74566;  
XX  
DT 15-JUL-2004 (first entry)  
XX  
DE Thale cress cDNA repressed in E2Fa/Dpa expressing plants SegID 2461.  
XX  
KW gene; ss; plant; transgenic; E2Fa/Dpa transcription factor;  
KW growth regulator; animal feed product; thale cress;  
KW cell wall biosynthesis; nitrogen metabolism; carbon metabolism.  
XX  
OS Arabidopsis thaliana.  
XX  
PN WO2004035798-A2.  
XX  
PD 29-APR-2004.  
XX  
PF 20-OCT-2003; 2003WO-BP011658.  
XX  
PR 18-OCT-2002; 2002EP-00079408.  
XX  
PA (CROP-) CROPPDESIGN NV.  
XX  
PI Inze D, De Veylder L, Vlieghe K;  
XX  
DR WPI; 2004-348466/32.  
XX  
DR P-PSDB; ADN74567.  
XX  
PT Altering plant characteristics, useful for producing plants for enzyme or  
PT pharmaceutical production comprises modifying in a plant, expression of  
PT one or more nucleic acids and/or modifying level or activity of one or  
PT more proteins.  
XX  
PS Claim 1; SEQ ID NO 2461; 134pp; English.  
XX  
CC This invention relates to a novel method for altering one or more plant  
CC characteristics. Specifically, it refers to identifying genes that are up  
CC - or down-regulated in transgenic plants overexpressing the heterodimeric  
CC E2Fa/Dpa transcription factor of Arabidopsis and using these sequences to  
CC alter plant characteristics accordingly. The present invention describes  
CC generating transgenic plants for the production of growth regulators,  
CC enzymes, therapeutics, pharmaceuticals and animal feed products, where  
CC the altered plant characteristics are selected from increased yield or  
CC biomass, enhanced survival capacity, stress tolerance, plant architecture  
CC or physiology, altered endoreduplication, biochemistry, signal  
CC transduction, storage lipid mobilisation and/or altered photosynthesis,

each relative to the corresponding wild type plants. Accordingly, these sequences can also be useful as positive or negative selectable markers during transformation of cells or tissues. The identified genes play a role in a variety of biological processes such as DNA replication, cell wall biosynthesis, nitrogen and/or carbon metabolism or they function as transcription factors. This polynucleotide sequence is thale cress cDNA represented 1.3 fold or more in plants overexpressing the E2ra/Dpa transcription factor, given in an exemplification of the invention.

Sequence 954 BP; 288 A; 198 C; 229 G; 239 T; 0 U; 0 Other;

**Alignment Scores:**

Pred. No.:	2.06e-153	length:	954
Score:	1611.00	Matches:	311
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	12	Gaps:	0

US-10-690-564-2 (1-317) X ADN74566 (1-954)

QY	1	MetLathrileuysValSeraspSerValProAlaProSerAspAspAlaGluGluLeu	20
Db	1	ATGGCGACTCTTAAAGTTTCGATTCGTCTCGCTCTTCGAATGAAGCGCAATTG	60
QY	21	ArgThrAlaPheGluGlyTTPGlyThrAsnGluAspLeuLeileSerIleLeuAlaHis	40
Db	61	AGAACCCCTTTTGAAGATGGGGTACGACAGGACTTATATCATCATCTTGGCTCAC	120
QY	41	ArgSerAlaGluGluArgValIleArgGluAlaThrIleGluThrTyrGlyGluAsp	60
Db	121	AGAAGTCTGAACAGAGAAAGTCATCAGGCAAGCATACCGAAACCTACGCGAAGAC	180
QY	61	LeuLeuIlePheThrLeuAspIleGluLeuSerAspAspPheGluAlaAlileuLeuTyr	80
Db	181	CTTCTCAAGACTCTTGCACAAAGAGCTCTCAAGATTTGAAAGAGCTATCTTGTGTGG	240
QY	81	ThrLeuGluIleProGlyGluArgAspAlaLeuLeuAlaAsnGluAlaThrIleAspGlyPhe	100
Db	241	ACTCTTGAACCCGGTGAAGCTGATGCTTTRTTGGCTTAAGAGCTACAAAAGATGAGCT	300
QY	101	SerSerAsnGluValLeuMetGluValAlaCysThrArgThrSerThrGluLeuLeuHis	120
Db	301	TCAAGCAACCAAGTCTTATGAAAGTTGGCTTGACAAAGACATCAACGAGCTGCTCAC	360
QY	121	AlaArgGluAlaThrHisAlaArgTyrIleValSerLeuGluGluAspValAlaHisHis	140
Db	361	GCTAGGGCAACCTTACCCATGCTGCTGCACAAAGAACTCTTGAAGAGAGAGTGGCTCACAC	420
QY	141	ThrThrGlyAspPheArgIleLeuLeuAlaSerLeuValThrSerTyrArgTyrGluGly	160
Db	421	ACTACCGGTACCTTCAGAAAGCTTTTGGTTTCTCTTGTTACCTCATCAGGTACGAGGA	480
QY	161	AspGluValAsnMetThrLeuAlaIleArgGluAlaIleValIleGluValIleGlyHis	180
Db	481	GATGAAGTGAACATGACATGGCTTGAACAAGAGCTAAGCTGCTCATGAGAAATCAAG	540
QY	181	AspIleHisIleTyrAsnAspGluAspValIleIleArgIleLeuSerThrArgSerIleAlaGln	200
Db	541	GACAAAGCATACAAATGATGAGATGATTATTAGATCTGTGTCCACAAGAAAGAAAGCTCAG	600
QY	201	IleAsnAlaThrPheAsnArgTyrGluAspAspHisGlyGluGluIleLeuIleSerLeu	220
Db	601	ATCATATCTTACCTTTAACCTTTCACAAAGATGATCATGGGAGAGAAATTTCTCAAGAGCTTT	660
QY	221	GluGluGluIleAspAspAspAspIlePheLeuAlaLeuLeuArgSerThrIleGlnCysLeu	240
Db	661	GAGAGAAAGAAATGATGATGACAAGTCTCTTGCACTTTGAAGTCAACCATTCAGAGCTTG	720
QY	241	ThrArgProGluLeuIleTyrPheValAspValLeuArgSerAlaIleAsnIleThrGlyThr	260
Db	721	ACAAGACCAAGCTTATCTTTGTGATAGTCTTCTTCGTTCAGCANATCAACAAAATCGGAAGT	780

QY	261	AspGluGlyValAlaLeuThrArgIleValThrThrArgAlaGluIleAspLeuLysValIle	260
Db	781	GATGAAGAGCACTCACTGAAATTGTGACCAACAAGACTGAGATTGACTGAAGTCAAT	840
QY	281	GlyGluGluIuTrGlnArgArgAsnSerIleProLeuGluLysAlaIleThrLysAspThr	300
Db	841	GGAGAGGAGTACCAAGCGCAGAAACAGCAATCTCTTGGAAAGCTATTACCAAGACACT	900
QY	301	ArgGlyAspTrpTrGluLysMetLeuValAlaLeuLeuGluGlyLysAspAla	317
Db	901	CGTGAGATTACAGAAAGATGCTCTGTGCACTTCGGTGAAGATGATGCT	951

## RE

ID AAC34418 standard; DNA; 1192 BP.

AAC34418

DT 17-OCT-2000 (first entry)

**DE Arabidopsis thaliana DNA fragment SEQ ID NO: 6580.**

KW Hybridisation assay; genetic mapping; gene expression control;

KW promoter; termination sequence; ss.

OS *Arabidopsis thaliana*.

PN EP1033405-A2.

PD 06-SEP-2000.

25-FEB-2000; 2000EP-00301439

PR 25-FEB-1999; 99US-0121825P

PR 09-MAR-1999; 99US-0123548P

PR 25-MAR-1999; 99US-0126264P

PR 01-APR-1999; 99US-0127462P

PR 08-APR-1999; 99US-0128714P

PR 19-APR-1999; 99US-0130077P

PR 23-APR-1999; 99US-0130510P

PR 28-APR-1999; 99US-0131449P

PR 30-APR-1999; 99US-0132407P

PR 05-MAY-1999; 99US-0132485P

PR 06-MAY-1999; 99US-0132487P

PR 11-MAY-1999; 99US-0134256P

PR 14-MAY-1999; 99US-0134219P

PR 14-MAY-1999; 99US-0134370P

PR 19-MAY-1999; 99US-0134941P

PR 21-MAY-1999; 99US-0135353P

PR 25-MAY-1999; 99US-0136021P

PR 28-MAY-1999; 99US-0136782P

PR 03-JUN-1999; 99US-0137528P

PR 07-JUN-1999; 99US-0137724P

PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.

PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157765P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161922P.  
PR 28-OCT-1999; 99US-0161933P.  
PR 29-OCT-1999; 99US-0162142P.

Alignment Scores:  
Pred. No. :  
Score:  
Percent Similarity:

2,74e-153  
1611.00  
100.00%

Length: 1192  
Matches: 317  
Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 3 Gaps: 0  
US-10-690-564-2 (1-317) x AAC34549 (1-1192)

QY 1 Mec1aThrLeuValSerAspSerValProAlaProSerAspAspAlaGluGluLeu 20  
Db 36 ATGGCGACTCTTAAGTTTCTGATTCCTGCTCTTCTGATGATGCGAGCAATTG 95  
QY 21 ArgThrAlaPheGluGlyTrpGlyThrAsnGluAspLeuIleIleSerIleLeuAlaHis 40  
Db 96 AGAACCCCTTTTGAAGATGGGTACAAACGAGACTTGCATATCAATCTTGCTGCAC 155  
QY 41 ArgSerAlaGluGluAspGlyValIleArgGlnAlaThrHisGluTrpTrpGlyGluAsp 60  
Db 156 AGAAGCTCTGAACGAGAAAGATGATCAGCGACATCAACATGAACTCAGGTGAAGAC 215  
QY 61 LeuLeuValThrLeuAspGlyGluLeuSerAspPheGluArgAlaIleLeuLeuTrp 80  
Db 216 CTTCTCAAGACTCTTGACAAAGACTCTCTACAGACTTCGAGAGACTTACTTGTGCG 275  
QY 81 ThrLeuGluProGlyGluArgAspAlaLeuLeuAlaAsnGluAlaThrIleArgTrpThr 100  
Db 276 ACTCTGAACCCGGTGAAGCTGATGCTTTATGCTATGAGACTTACAAAAGATGACT 335  
QY 101 SerSerAsnGlnValLeuMetGluValAlaCysThrArgThrSerThrGlnLeuLeuHis 120  
Db 336 TCAAGCAACCAAGCTTATGAGAGTGTGCAAGAGCAATCAACGACGCTGCTTAC 395  
QY 121 AlaArgGlnAlaThrHisAlaArgTrpIleLeuSerLeuGluAspValAlaHisHis 140  
Db 396 GCGAGGCAAGCTTACATGCTCGCTCAAGAACTCTTGAAGAGCGTGGCTCACCAC 455  
QY 141 ThrThrGlyAspPheArgGlyLeuLeuValSerLeuValThrSerTrpArgTrpGlyGly 160  
Db 456 ACTACCGGTACTTCAAAAGCTTTGTCTCTCTTACTATCAAGTACAGTACGAAGA 515  
QY 161 AspGluValAsnMetThrLeuAlaIleGluGlnAlaIleGluValHisGluValIleLeu 180  
Db 516 GATGAACTGAAACATGACATTTGGCTACAGCAAGACTAAGTGGTCCATGAGAAATCAG 575  
QY 181 AspIleHisTrpAsnAspGluAspValIleArgIleLeuSerThrArgSerIleValGln 200  
Db 576 GACAAGACTACATGATGATGAGATGATTTAGATCTTGTCCACAAAGCAAGCTCAG 635  
QY 201 IleAsnAlaThrPheAsnArgTrpGlnAspAspHisGlyGluGluIleLeuSerLeu 220  
Db 636 ATCAATGCTACTTTTAAACGTTACCAAGATGATCATGGCGAGAAATTTCTCAAGACTT 695  
QY 221 GluGluGlyAspAspAspAspAspPheLeuAlaLeuLeuArgSerThrIleGlnCysLeu 240  
Db 696 GAGAAAGAGATGATGATGACAAAGTCTTTCATCTTTGAGTCAACATTCAGTGGCTTG 755  
QY 241 ThrArgProGluLeuValTrpPheValAspValLeuArgSerAlaIleAsnValThrGlyThr 260  
Db 756 ACAAGACCAAGACTTTCTTTGTCGATGTTCTTGTTCAGCAATCAACAAACTGGAACT 815  
QY 261 AspGluGlyAlaLeuThrArgIleValThrThrArgAlaGluIleAspLeuValIle 280  
Db 816 GATGAAGAGACACTCATTAGATTGTGACCAAGACTCAGAAFTGACTTGAAAGTCAAT 875  
QY 281 GlyGluGlyTrpGlnArgAsnSerIleProLeuGluValAlaIleThrIleAspThr 300  
Db 876 GGAAGAGAGTACCAAGCGCAGAAACAGATTCCTTTGAGAAAGTATTATCAAGACACT 935  
QY 301 ArgGlyAspTrpGlyIleMetLeuValAlaLeuLeuGlyGluAspAspAla 317  
Db 936 CGTGAAGATTACGAAAGATGCTGCTCGCACTTCTCGTGAAGATGATGCT 986

RESULT 4  
AAC34549  
ID AAC34549 standard; DNA; 1212 BP.

XX AAC34549;  
AC 17-OCT-2000 (first entry)  
DT  
XX  
XX  
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 7037.  
XX  
XX Hybridisation assay; genetic mapping; gene expression control;  
KW protein identification; signal transduction pathway; metabolic pathway;  
XX promoter; termination sequence; ss.  
XX Arabidopsis thaliana.  
PN EP1033405-A2.  
PD 06-SEP-2000.  
XX  
XX  
PF 25-FEB-2000; 2000EP-00301439.  
XX  
XX 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 28-APR-1999; 99US-0130891P.  
PR 30-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 04-MAY-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.

PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 24-JUN-1999; 99US-0140354P.  
PR 28-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 28-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.

PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 30-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151033P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161922P.  
PR 28-OCT-1999; 99US-0161993P.  
PR 29-OCT-1999; 99US-0162142P.

## Alignment Scores:

Pred. No.: 2-8e-153 Length: 1212  
Score: 1611.00 Matches: 317  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 3 Gaps: 0

US-10-690-564-2 (1-317) x AAG34549 (1-1212)

QY 1 MetAlaThrLeuIyValIserAspSerVaIProAlaProSerAspAspAlaGluGluLeu 20  
DB 32 ATGGGACTCTTAAAGATTTCGATTCGTCTCGCTCTTCGATGAGCTGAGCAATTG 91  
QY 21 ArgThrAlaPheGluGlyTTrpGlyThrAsnGluAspLeuIleIleSerIleLeuAlaHis 40  
DB 92 AGAACCGCTTTTGAAGAGATGGGGTCAGACGAGGACTTGATCATATCTTGCTGCAC 151  
QY 41 ArgSerAlaGluGlnArgIyValIleArgIlnalATyRHsgluThrTyroIyGluAsp 60

```
Db 152 AGAGTCTGAACAGAGGAAGCATCATGACGACACATACCTGAACCTGAGGGAAGAC 211
Qy 61 LeuLeuYThrThrLeuAspLysGluLeuSerAsnAspPheGluArgAlaLeuLeuTyr 80
Db 212 CTTCTCAAGACTCTTGACACAGAGACTCTCTAACCACTTCGAGAGAGCTATCTGTGTGG 271
Qy 81 ThrLeuGluProGlyGluArgAspAlaLeuLeuAlaGluAlaThrThrArgTyrThr 100
Db 272 ACTCTTAACCCGGTGAAGCGTGAAGCTTATGTGCTATGAGTACAAAAGATGACAT 331
Qy 101 SerSerAsnGluValLeuMetGluValAlaCysThrArgThrSerThrGluLeuHis 120
Db 332 TCAGCAACCAAGTCTTATGAGAGTCTTGACACAGACATCAACGCACTCTTCAC 391
Qy 121 AlaArgGlnAlaTyrHisAlaArgTyrLysSerLeuGluAspValAlaHisHis 140
Db 392 GCTAGGCAAGCTTACCATGCTCGCTACAGAAAGTCTTGAAAGAGAGCTGCTCACAC 451
Qy 141 ThrThrGluAspPheArgLysLeuLeuValSerLeuValThrSerTyrArgTyrGluGly 160
Db 452 ACTACCGGTGACTTCAGAAAGCTTTGGTTCCTGTTACTCATACAGGTACGAAGA 511
Qy 161 AspGluValAsnMetThrLeuAlaLysGlnGluAlaLysLeuValHisGluLysIleLys 180
Db 512 GATGAAGTGAACATGACATTTGGCTAGACAGAACTAAGCTGCTCCATGAAAAATCAG 571
Qy 181 AspLysHisTyrAsnAspGluAspValIleArgIleLeuSerThrArgSerLysAlaGln 200
Db 572 GACAAAGCACTACATGATGAGAGATGTTATGAAATCTTGTCACAAAGCAAGCTCAG 631
Qy 201 IleAsnAlaThrPheAsnArgTyrGluAspAspHisGlyGluGluIleLeuLysSerLeu 220
Db 632 ATCATATCTACTTTAAACCGTTACCAAGATGATCATGCGAGAAATTCACAAAGCTCT 691
Qy 221 GluGluGluLysAspAspAspAspLysPheLeuAlaLeuLeuArgSerThrIleGlnCysLeu 240
Db 692 GAGCAAGAGATGATGATGACAAAGTCTTCCTGTTGAGGTACACCATTCAGTGGCTTG 751
Qy 241 ThrArgProGluLeuTyrPheValAspValLeuArgSerAlaIleAsnLysThrGlyThr 260
Db 752 ACAAGACCAAGGCTTACTTGTGATGTTCTGCTGCTGCAAGATCAACAAAGCTGAAC 811
Qy 261 AspGluGluLysAlaLeuThrArgIleValThrThrArgAlaGluIleAspLeuLysValIle 280
Db 812 GATGAAGAGACACTCACTAGATTGTCAACCAAGAGCTGAGATTGACTTGAAAGTCAAT 871
Qy 281 GlyGluGluTyrGlnArgAsnSerIleProLeuGluLysAlaIleThrLysAspThr 300
Db 872 GGAGAGAGATGCCAGCCGACAGACAGACATTCCTTTGAGAAAGCCATTACCAAGACACA 931
Qy 301 ArgGlyAspTyrGluLysMetLeuValAlaLeuLeuGlyGluAspAspAla 317
Db 932 CGTGAAGATTACGAGAAAGATGCTGCTGCACTTCGCTGTAAGATGATGCT 982
RESULT 5
AAC33439
ID AAC33439 standard; DNA, 1182 BP.
XX
XX AAC33439;
XX
DT 17-OCT-2000 (first entry)
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 3039.
XX
XX Hybridisation assay; genetic mapping; gene expression control;
KM protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX
XX Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
```

```
PD 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-00301439.
XX
XX 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126284P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 21-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140053P.
PR 23-JUN-1999; 99US-0140054P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-014187P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
```



PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144351P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144864P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145234P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 03-AUG-1999; 99US-0146389P.  
PR 04-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 05-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.

PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0159848P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161922P.  
PR 28-OCT-1999; 99US-0161933P.  
PR 29-OCT-1999; 99US-0162142P.

## Alignment Scores:

Pred. No.: 9 74e-149 Length: 1182  
Score: 1566.00 Matches: 314  
Percent Similarity: 98.74% Conservative: 0  
Best Local Similarity: 98.74% Mismatches: 1  
Query Match: 97.21% Indels: 3  
DB: 3 Gaps: 1

US-10-690-564-2 (1-317) x AAC33439 (1-1182)

QY 1 MetAlaThrLeuLysValSerAspSerValProAlaProSerAspAlaGluGluLeu 20  
DB 33 ATGGGACCTCTTAAGGTTTCTGATTCGTCTGCTCCTTGAGATGCTGACGATTG 92  
QY 21 ArgThrAlaPheGluGly-TyrGlyThrAsnGluAspLeuIleIleSerIleLeuLahI 40  
DB 93 AGAAGCGCT-----TTGATGGGGTACAACAGGACCTTGACATCATCTTGCTCA 146  
QY 40 sArgSerAlaGluGlnIleArgLysValIleArgGlnAlaTyrRhsGluThrTyrGlyGluAs 60  
DB 147 CAGAAAGTCTGAACGAGGAAAGTAGTACGGCAAGCATACCTGAAAGGAGTGA 206  
QY 60 pleuLeuLysThrLeuAspLysGluLeuSerAsnAspPheGluArgAlaIleLeuLeuTr 80  
DB 207 CCTTCTCAAGACTCTTGGAAGAGACTCTTAACGACTTCGAGAGAGCTATCTGTGTG 266  
QY 80 pThrLeuGluProGlyGluArgAspAlaLeuLeuAlaAsnGluAlaThrLysArgTyrPth 100  
DB 267 GACTCTTGAAACCGGAGACGTGATGCTTATTTGGCTAATGAAGCTACAAAAGATGAC 326  
QY 100 rSerSerAsnGlnValLeuMetGluValAlaGlyThrArgThrSerThrGluLeuLeuH 120  
DB 327 TTCAGCAACCAAGTGCCTTATGGAAGTAGCTTGCAAGGAGATCATCAAGAGCTGCTTA 386

```

QY 120 sAlaArgGlnAlaTyrHisAlaArgTyrLysSerLeuGluGluAspValAlaHisH 140
DB 387 CCGAGGAGCAAGCTTACATGCTCGCTACAGAAATCTCTTGAAGAGAGAGCTGGCTCA 446
QY 140 gThrThgTlYAspPheArgLysLeuLeuValSerLeuValThrSerTyrArgTyrGlu 160
DB 447 CACTACCGGTGACTTCAGAAAGCTTTGGTCTCTCTTGTACTTACCTACAGAGTACGA 506
QY 160 yAspGluValAspMetThrLeuAlaLysGlnGluAlaLysLeuValHisGluLysIle 180
DB 507 AGATGAAGTAAACATGATGCTTGGTAAAGAGAAAGTAACTGGTTCATAGAAACAA 566
QY 180 sAspLysHisTyrAspAspGluAspValIleArgIleLeuSerThrArgSerLysAla 200
DB 567 GGACACGACTTACATATATAGATGTATTAAGATCTTGTCCACAGAAAGCAAGCTCA 626
QY 200 nIleAsnAlaThrPheAsnArgTyrGlnAspAspHisGluGluGluLysSerLe 220
DB 627 GATCAATGACTACTTTTAAACGCTTACCAAGATGATCATGGGAGGAATTCCTCAAGAG 686
QY 220 uGluGluGluYAspAspAspAspLysPheLeuAlaLeuLeuArgSerThrIleGln 240
DB 687 TGAGGAGAGAGATGATGATACAAAGTCTTGTGACCTTTGAGGTCACATTCAGTGTCT 746
QY 240 uThrArgProGluLeuTyrPheValAlaAspValIleuArgSerAlaIleAsnLysTh 260
DB 747 GACAAAGCCAGAGCTTACTTGTGATGATCTTGTGACGATTCACAAACAACTGGAAC 806
QY 260 rAspGluGluYAlaLeuThrArgIleValIleThrThrArgAlaGluIleAspLeu 280
DB 807 TGATGAAGAGAGCACTCACTGAAATTTGACACAAAGCTGAGATTGACTTGAAGGTCA 866
QY 280 eGluGluGluTyrGlnArgArgAsnSerIleProLeuGluLysAlaIleThrLys 300
DB 867 TGAGAGGAGAGTACAGGAGGAGAAACAGCATTTCTTGGAGAAAGCTATTACCAAGAC 926
QY 300 rAspGluYAspTyrGluLysMetLeuValAlaLeuLeuGluGluAspAspAla 317
DB 927 TCGTGAGATTACGAGAAAGATGCTCGTGGCACTTCTGTGAGAGATGATGCT 978

```

RESULT 6  
ADR63461  
ID ADR63461 standard; cDNA, 1184 BP.

AC ADR63461;

DT 02-DEC-2004 (first entry)

DE Cotton cDNA sequence, SEQ ID 4242.

XX Cotton; ss; plant; cold tolerance; growth rate; cell cycle pathway;  
XX drought tolerance; plant disease resistance; galactomannan; lignin;  
XX plant growth regulator; heat tolerance; herbicide tolerance;  
XX homologous recombination; extreme osmotic condition tolerance;  
XX pathogen resistance; pest resistance; yield; photosynthesis; seed oil;  
XX stress resistance.

OS *Gossypium hirsutum*.

XX US2004181830-A1.

XX 16-SEP-2004.

XX 29-JAN-2004; 2004US-00767795.

XX 07-MAY-2001; 2001US-00849529.

XX 12-DEC-2001; 2001US-00021323.

XX (KOVA/) KOVALIC D K.

XX (ZHOU/) ZHOU Y.

XX (CAOY/) CAO Y.

XX Kovalic DK, Zhou Y, Cao Y;

```

XX XX
DR WPI; 2004-667718/65.
XX XX
PT New recombinant nucleic acid molecules and polypeptides from Gossypium
PT hirsutum, useful for producing plants with improved biological
PT characteristics (e.g. improved plant cold or drought tolerance).
XX XX
PS Claim 1; SEQ ID NO 4242; 14pp; English.
XX XX
CC The invention relates to a recombinant polynucleotide comprising any of
CC the 58798 Cotton plant cDNA sequences mentioned in the specification.
CC Also a recombinant polypeptide comprising any of the 58798 amino acid
CC sequences mentioned in the specification and producing a plant having an
CC improved property. Producing a plant having an improved property
CC comprises transforming a plant with a recombinant construct comprising a
CC promoter region functional in a plant cell operably joined to a
CC polynucleotide comprising a coding sequence for a polypeptide associated
CC with the property, and growing the transformed plant. The polypeptide is
CC useful for improving plant cold tolerance, manipulating growth rate in
CC plant cells by modification of the cell cycle pathway, improving plant
CC drought tolerance, providing increased resistance to plant disease,
CC producing galactomannan (or lignin or plant growth regulators), increasing
CC plant heat tolerance, improving plant tolerance to herbicides, increasing
CC the rate of homologous recombination in plants, improving plant tolerance
CC to extreme osmotic conditions or to pathogens or pests, improving yield
CC by modification of photosynthesis, modifying seed oil or protein yield
CC and/or content, improving yield by modification of carbohydrate, nitrogen
CC or phosphorus use and/or uptake, or improving yield by providing improved
CC plant growth and development under at least one stress condition. The
CC polynucleotide and polypeptide may also be used in recombinant DNA
CC constructs, in physical arrays of molecules, as plant breeding markers,
CC or in computer-based storage and analysis systems. The present sequence
CC is a Cotton plant cDNA of the invention. NOTE: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?docID=20040181830. However only 6585
CC polynucleotide sequences were available, the remaining 52213
CC polynucleotides and all 58798 protein sequences were not present.
XX XX
SQ Sequence 1184 BP; 375 A; 233 C; 281 G; 295 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 6 07e-112 Length: 1184
Score: 1202.50 Matches: 230
Percent Similarity: 85.30% Conservative: 37
Best Local Similarity: 73.48% Mismatches: 45
Query Match: 74.64% Indels: 1
DB: 13 Gaps: 1

US-10-690-564-2 (1-317) x ADR63461 (1-1184)
QY 1 MetAlaThrLeuLysValSerAspSerValProAlaProSerAspAspAlaGluGlnLeu 20
DB 70 ATGGCCACTCTTAACAGTGCACAGACAGTTCCTCGGTGTCTGAAATGTGAAACGCTA 129
QY 21 ArgThrAlaPheGluGluTyrTrpGlyThrAsnGluAspLeuIleIleSerIleLeuAlaHis 40
DB 130 AGAAAGCCCTTTCAGAGATGGGAACTAATGAGGCGCTTAATCATATATATTGGGTGAC 189
QY 41 ArgSerAlaGluGluLysValIleArgGlnAlaTyrHisGluThrTyrGluLysAsp 60
DB 190 AGAAATGCCAGACAGCAAGAACTTGATTCAGAAACCTTACGCTGAACCTATGAGAGAGAT 249
QY 61 LeuLeuLysThrLeuAspLysGluLeuSerAspAspPheGluLysAlaIleLeuLeuTyr 80
DB 250 CTCCTCAAGCACTAGACAAAGAGCTCTGAAATGACTTTGAGAGGCTGTTTGGCTTTGG 309
QY 81 ThrLeuGluProGluYAlaArgAspAlaLeuLeuAlaAsnGluAlaThrLysArgTyrThr 100
DB 310 GCTCTTGATCTCGTGAAAGTATGCTCTTTGGCTTAATGAAGCAACCAAAAGGTGGACT 369
QY 101 SerSerAsnGluValLeuMetGluValAlaCysThrArgThrSerThrGlnLeuLeuHis 120

```



[illegible]

```
Db      932 AGTGAAGATACGAGAGGAAATAGCATCCCTCGATCGTGCATTTGCAAGACACT 991
Qy      301 ArgGlyAspTyrGluIysMetLeuValAlaLeuLeuGlyGluAspAspAla 317
Db      992 ACTGAGACATATGAAAATAATGCTTCTGGCATTTGCGACGCTGAGGCT 1042

RESULT 9
AAFA4756 standard; cDNA, 1293 BP.
XX
AC   AAF44756;
XX
DT   27-MAR-2001 (first entry)
XX
DE   Annexin-like protein coding sequence #1.
XX
KW   Cell death modulator; programmed cell death; PCD; apoptosis;
KM   forestry plant; ss.
XX
OS   Pinus radiata.
XX
PN   WO200075331-A1.
XX
PD   14-DEC-2000.
XX
PF   02-JUN-2000; 2000MO-NZ000086.
XX
PR   04-JUN-1999; 99US-00325932.
XX
PA   (GENE-) GENESIS RES & DEV CORP LTD.
PA   (FLET-) FLETCHER CHALLENGE FORESTS LTD.
PI   Flinn B, Lasham A;
XX
DR   WPI; 2001-061724/07.
DR   P-PSDB; AAB5730.
XX
PT   Novel defender against cell death polynucleotide useful for modulating
PT   programmed cell death pathway and specific development pathways in
PT   forestry plant.
XX
PS   Claim 1; Page 54-55; 142pp; English.
XX
CC   The present invention relates to coding sequences (see AAF44740-P44840
CC   and AAF44843-P44844) and proteins (see AAB5714-B5814) involved in
CC   programmed cell death (PCD; apoptosis). The coding sequences and proteins
CC   of the present invention are useful for modulating a PCD or cell death
CC   pathway and various developmental pathways in a forestry plant, by stably
CC   incorporating one of the present coding sequences into the genome of the
CC   forestry plant, where the coding sequence provides a PCD pathway that is
CC   not present in a native form of the forestry plant
XX
SQ      Sequence 1293 BP; 359 A; 290 C; 313 G; 331 T; 0 U; 0 Other;

Alignment Scores!
Pred. No.:      9.34e-109      Length:      1293
Score:          1171.50      Matches:      226
Percent Similarity: 83.60%      Conservative: 39
Best Local Similarity: 71.29%      Mismatches: 51
Query Match:    72.72%      Indels:      1
DB:            4      Gaps:      1

US-10-690-564-2 (1-317) x AAF44756 (1-1293)

Qy      1 MetAlaThrLeuLysValSerAspSerValProAlaProSerAspAspAlaGluGluLeu 20
Db      95 ATGTCTCACTCTCAACCTCCGCGACGACCTGCTGTAGCCGATACCTGCGAGCAGCTC 154
Qy      21 ArgThrAlaPheGluGlyTyrGlyThrArgGluAspLeuLeuIleSerIleLeuAlaHis 40
Db      155 CGAGACGCTTCCACGAGATGGGGAACAAATGAACACTGATTCATATTCATATTTGGGTCA 214
Qy      41 ArgSerAlaGluGluArgLysValIleArgGlnAlaTyrHisGluThrTyrGlyGluAsp 60
```

```
Db      215 AGCAATGCGCGCCAGAGAGAAAGCTGATTCGGCAAACTATGCCGAGACTTACGGCAGAGAC 274
Qy      61 LeuLeuLysThrLeuAspLysGluLeuSerAsnAspPheGluArgAlaIleLeuLeuTyr 80
Db      275 CTCTCAAGGCAATTTGGACAGAAAGCTTACCAATGATTTTCGAGAGGCTGGTCTCTTTGG 334
Qy      81 ThrLeuGluProGlyGluArgAspAlaLeuLeuAlaAsnGluAlaThrLysArgTyrThr 100
Db      335 TCACCTTGATCCGGCTGAAAGCTGATCGTACTTGGCGAATGAAAGCAAGAAAGATGACT 394
Qy      101 SerSerAsnGluValLeuMetGluValAlaCysThrArgThrSerThrGluLeuLeuHis 120
Db      395 TCAGCAACACCAAGTCTCTCATGGAATAGCTTCACAGAGCTCCGACAGACTGCTTATG 454
Qy      121 AlaArgGlnAlaTyrHisAlaArgTyrLysSerLeuGluGluAspValAlaHisHis 140
Db      455 GCAGACACAGCATATCATGATCCGATACAGAAAGTAAATGAGAGAGAGCTGCTACAC 514
Qy      141 ThrThrGlyAspPheArgLysLeuLeuValSerLeuValThrSerTyrArgTyrGluGly 160
Db      515 ACACTGAGATTTTCTGTAAGTCTGCTGTAACCTCTTGGAGAGCTCTACCGTAAATATGCA 574
Qy      161 AspGluValAsnMetThrLeuAlaLysGlnGluAlaLysLeuValHisGluLysIleLys 180
Db      575 GATGAGGTGAATATATGACTTTGGCAAAAGAGGCTTAAGATATCTCACAGAAAGATCTCA 634
Qy      181 AspLysHisIleTyrAsnAspGluAspValIleArgIleLeuSerThrArgSerIleValGln 200
Db      635 GAGAAAGGCTTATGAGCCATGAGATGCTCATAGGATTTGGCTATTCAGAGACAAACACAG 694
Qy      201 IleAsnAlaThrPheAsnArgTyrGlnAspAspHisGlyGluGluIleLeuLysSerLeu 220
Db      695 GTCAATGCTACCGCTGATCATCAAAATAGATTTGGAATGATATCAACAGGATTTG 754
Qy      221 GlnGluGlyAspAspAspLysPheLeuAlaLeuLeuArgSerThrIleGlnCysLeu 240
Db      755 AAAAAGCT--GATCCAAAAGACCGCTTCTCTACTATATGAGAGCTACAGTAAAGGCGCTG 811
Qy      241 ThrArgProGluLeuTyrPheValAspValLeuArgSerAlaIleAsnLysThrLysThr 260
Db      812 ACTGCCCTCGAAGATATTTTGAAGAAAGTCTCTCTCTAGCCATCAATTAAGGAGAAACA 871
Qy      261 AspGluGlyAlaLeuThrArgIleValThrThrArgAlaGluIleAspLeuLysValIle 280
Db      872 GATGAAGGGGCTCTGACAGAGATTTGCTACAGAGGCGCGAGTTGACATGAAGTTTATA 931
Qy      281 GlnGluGlyTyrGlnArgArgAsnSerIleProLeuGluLysAlaIleThrLysAspThr 300
Db      932 ACTGAGGAGTACCAAGAGGAGAAATGATCCCTTCGATCGGCGCATTTGTCAAGGACACT 991
Qy      301 ArgGlyAspTyrGluLysMetLeuValAlaLeuLeuGlyGluAspAspAla 317
Db      992 ACTGAGACATATGAAAATAATGCTTCTGGCATTTGCGACGCTGAGGCT 1042

RESULT 10
ADB94943 standard; cDNA, 1293 BP.
ID      ADB94943
XX
AC      ADB94943;
XX
DT      04-DEC-2003 (first entry)
XX
DE      Programmed cell death pathway protein annexin cDNA #1.
XX
KW      programmed cell death; plant development; plant cell cycle; ATL2; DAD1;
KW      Dnaase; 118; 1ad1; muellin-like aspartic protease; annexin; prohibitin;
KW      fen-like protein; rac2; retinoblastoma-related protein; SIN3; TRIP;
KW      TEGT; xylogenic Rnaase; pur-alpha; cyteine protease; RPP5-like protein;
KW      gp 91 NADPH oxidase subunit; NPR-like protein; BAG-1;
KW      defender against cell death; lethal leaf spot; lesion stimulating death;
KW      seven in absentia; transcription initiation factor;
KW      testis enhanced gene transcript; gene; ss.
```



PT New isolated polynucleotide useful for modulating programmed cell death,  
 PT altering the development cycle of plant cells, and subsequently modifying  
 PT plant development.

XX Claim 1; Page 138; 21app; English.

XX The invention describes an isolated polynucleotide (1) comprising a  
 CC nucleotide sequence that is one of 145 fully defined sequences of 221-  
 CC 3415 base pairs (bp), given in the specification, and/or its complements,  
 CC reverse complements, or sequences having 75, 90 or 95  
 CC % sequence identity to or that hybridise under stringent hybridisation  
 CC conditions to one of the 145 sequences. The methods and compositions of  
 CC the present invention to do with (1) are useful for modulating programmed  
 CC cell death and thereby altering the development cycle of plant cells, and  
 CC altering plant development. This sequence encodes a protein associated  
 CC with the programmed cell death pathway.

XX Sequence 1058 BP; 306 A; 227 C; 274 G; 251 T; 0 U; 0 Other;

#### Alignment Scores:

Pred. No.:	6,39e-105	Length:	1058
Score:	1132.50	Matches:	218
Percent Similarity:	83.91%	Conservative:	48
Best Local Similarity:	68.77%	Mismatches:	50
Query Match:	70.30%	Indels:	1
DB:	10	Gaps:	1

US-10-690-564-2 (1-317) x ADB94854 (1-1058)

QY 1 MetAlaThrLeuLysValSerAspSerValProAlaProSerAspAspAlaGluGlnLeu 20  
 Db 72 ATGGCGACATCCGATGCGCACCGTCGGTCCGTCGCCGCTGAGATGCCGAGCGTC 131  
 QY 21 ArgThrAlaIleGluGlyTTPGlyThrAsnGluAspLeuIleIleSerIleLeuAlaHis 40  
 Db 132 CAAAGAGCTTCCGAGATGGGAGACGAAGAAATCTGATCATCCATACCTACGCTCAC 191  
 QY 41 ArgSerAlaGluGlnArgLysValIleArgGlnAlaTyrHisGluThrTyrGlyGluAsp 60  
 Db 192 AGAAACGACCGCGACCGGAAAGTAAATCCGAAACATATCCGAGACATATGGGAAAGAT 251  
 QY 61 LeuLeuLysThrLeuAspLysGluLeuSerAspAspPheGluArgAlaIleLeuLeuTyr 80  
 Db 252 CTTCTCAAGCGCTGACGAAGAACTCTAGTGAAGATCTTGAGAGATCTTGCTGTGG 311  
 QY 81 ThrLeuGluProGlyGluArgAspAlaLeuLeuAlaAsnGlnAlaThrLysArgTyrThr 100  
 Db 312 ACCCTGATCTCGGAGCGTATGATCTTGTCCAAATGAAGCTACCAAGAGATTGACT 371  
 QY 101 SerSerAsnGlnValLeuMetGluValAlaCysThrArgThrSerThrGlnLeuLeuHis 120  
 Db 372 TCAGACACATGGGTTCTTCATGAAATGCTTGACACAGAGCTTCAATGAGATTATTCAG 431  
 QY 121 AlaArgGlnAlaTyrHisAlaArgTyrLysSerLeuGluGlnAspValAlaHisHis 140  
 Db 432 GTGAGCGACGCTATCATGCTCGTTAAAGAAATCTTGAAGAAACATCCGATATCAC 491  
 QY 141 ThrThrGlyAspPheArgLysLeuLeuValSerLeuValThrSerTyrArgTyrGluGly 160  
 Db 492 ACTACTGGGATTTCCGCAAGCTGCTTGTCTTGCAAGATGACCTTCGATATGAGGG 551  
 QY 161 AspGluValAsnMetThrLeuAlaLysGlnGluAlaLysLeuValHisGluLysIleLys 180  
 Db 552 CCGAGAGTGAACATGACATTTGGCGAGATCAGAGGCTAAAGTACTTCAATGAAAGATTCC 611  
 QY 181 AspLysHisTyrAsnAspGluAspValIleArgIleLeuSerThrArgSerLysAlaGln 200  
 Db 612 GAGAAAGCTTACATCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 671  
 QY 201 IleAsnAlaThrPheAsnArgTyrGlnAspAspHisGlyGluGlnIleLeuLysSerLeu 220  
 Db 672 CTTAATGCAACCTCAATTAATCAACAAATGAGTTGGGATGCGCATCAACAGAGATTCTG 731

QY 221 GluGluGlyAspAspAspLysPheLeuAlaLeuLeuArgSerThrIleGlnCysLeu 240  
 Db 732 ---MAGCGTANCCAAATGATGAATTTCTGAACTCGTGAAGTATGCAATTAAGCGCTTG 788  
 QY 241 ThrArgProGluLeuTyrPheValAspValLeuArgSerAlaIleAsnLysThrGlyThr 260  
 Db 789 ACTTATCCAGAGAAATGACTTTCAGAAAGTACTGCTCGGCATCAACAGCTGGGAACA 848  
 QY 261 AspGluGlyAlaLeuThrArgIleValThrThrArgAlaGluIleAspLeuValIle 280  
 Db 849 GATGAATGGGCTTTCACAGAGTACTCACACTCGTGGGAAAGTGAACATGACGCGCATC 908  
 QY 281 GlyGluGlyTyrGlnArgArgAsnSerIleProLeuGluLysAlaIleThrLysAspThr 300  
 Db 909 AAAGAAAGTACCAACAAAGAAAGAAAGTGTCTTCTTCAATGCGGCTATTGCAATGACACA 968  
 QY 301 ArgGlyAspTyrGluLysMetLeuValAlaLeuLeuGlyGluAspAspAla 317  
 Db 969 TCTGGGGACTACAAAGAGATGCTCTCGCTTGTGTGACATGAGAGATGCT 1019

RESULT 12  
 ADF37996  
 ID ADF37996 standard; cDNA; 1157 BP.  
 AC ADF37996;  
 DT 12-FEB-2004 (first entry)  
 XX Synchronised tobacco BY2 cDNA sequence SEQ ID NO:55.  
 DE Identification; validation; plant; agrochemical; gene; ss.  
 OS Nicotiana tabacum.  
 EN WO2003085115-A2.  
 PD 16-OCT-2003.  
 PF 08-APR-2003; 2003WO-BP003703.  
 XX 10-APR-2002; 2002EP-00447062.  
 PR 15-JUL-2002; 2002US-0396124P.  
 XX (CROP-) CROPDISEIGN NV.  
 PA Inze D, Broekaert W;  
 PI WPI, 2003-804308/75.  
 DR Identifying and validating plant genes or proteins as targets for  
 PT agrochemicals, useful for producing agrochemical-resistant plants,  
 PT comprises determining and down regulating the gene or protein expression  
 PT profiles of a plant.  
 XX Claim 12; SEQ ID NO 55; 183bp; English.  
 PS The present invention describes a method for identifying and validating  
 CC plant genes/proteins as targets for agrochemicals comprising determining  
 CC the gene or protein expression profiles of a plant and downregulating the  
 CC expression of the gene or protein in the plant or plant cell. Also  
 CC described: (1) methods for screening candidate agrochemical compounds,  
 CC comprising the use of the above method or the use of any of the 785 fully  
 CC defined nucleotide sequences (ADF37942 to ADF38726) or protein sequences,  
 CC or their homologues, functional fragments or derivatives; (2) a method  
 CC for producing an agrochemical resistant plant, comprising the use of the  
 CC above-mentioned nucleotide or protein sequences; (3) an isolated nucleic  
 CC acid that is identified by any of the above methods or that comprises at  
 CC least a part of a nucleic acid sequence chosen from any of the 785  
 CC nucleotide sequences given in the specification; (4) a plant tolerant to  
 CC an agrochemical, in which the expression level of one or more of the  
 CC nucleic acid sequences given in the specification is modulated; and (5) a  
 CC harvestable part of the plant described above. The method is useful in  
 CC identifying and validating plant targets for agrochemicals or in

CC producing an agrochemical resistant plant. The nucleic acid or protein  
 CC can be used as a target for an agrochemical compound, particularly  
 CC herbicide. The present sequence represents a synchronised tobacco BY2  
 CC cDNA nucleotide sequence which is used in the exemplification of the  
 CC present invention.

XX  
 SO Sequence 1157 BP; 362 A; 213 C; 270 G; 312 T; 0 U; 0 Other;

# Alignment Scores:

Pred. No.:	3.07e-102	Length:	1157
Score:	1106.50	Matches:	214
Percent Similarity:	81.07%	Conservative:	43
Best Local Similarity:	67.51%	Mismatches:	59
Query Match:	68.68%	Indels:	1
DB:	10	Gaps:	1

US-10-690-564-2 (1-317) x ADF37996 (1-1157)

```

QY 1 MetAlaThrLeuValSerAspSerValProAlaProSerAspAspAlaGluGluLeu 20
DB 45 ATGGCGAGCTTTAAAGTCCCAACATCTGTTCCAGAACCTTATGAGAGTGGCAACATC 104
QY 21 ArgThrAlaPheGluGlyTyrGlyThrAsnGluAspLeuLeuSerIleLeuAlaHis 40
DB 105 AAAAAGCTTTTGGCTGATGGGGGTACAAATGAGGCACTTATTATTCAGATTCTGGCAT 164
QY 41 ArgSerAlaGluGlnAlaGlyValIleArgGlnAlaTyrHisGluThrTyrGlyGluAsp 60
DB 165 AGAATGACACACACCAACCAAGTTATCCGAAACCTTATCTGCACCTTATGAGAGAGAT 224
QY 61 LeuLeuValThrLeuAspValGluLeuSerAsnAspPheGluAspAlaIleLeuLeuTyr 80
DB 225 CTCTCTCAAGACCTTGAGTGGTGAACCTGACCAAGTATTTTCAAGCTGGACGCTTCTGG 284
QY 81 ThrLeuGluProGlyGluArgAspAlaLeuLeuAlaAsnGlnAlaThrLysArgTyrThr 100
DB 285 ACTTTGAGTCTGCTGGAGCGGACGCTACTTGGTTATGAGAGTACCAACGCTGACT 344
QY 101 SerSerAsnGlnValLeuMetGluValAlaCysThrArgThrSerThrGlnLeuLeuHis 120
DB 345 TCTAGCAATGGGTATCTTGGAAATGCTGTGTACAAAGCTTCTGTGATGATCTTTAAG 404
QY 121 AlaArgGlnAlaTyrHisAlaArgTyrLysSerLeuGluGluAspValAlaHisHis 140
DB 405 GCGAGGAGGCTTACACATGCTCGATACAAAGAACCTTAAAGAAAGTGTGCTTATCAC 464
QY 141 ThrThrGlyAspPheArgLysLeuLeuValSerLeuValThrSerTyrArgTyrGluGly 160
DB 465 ACAACTGGGATTTCCGTAAGCTTTTGGTTCCTCTTTAATCGCATTCAGATACGAAAGCA 524
QY 161 AspGluValAsnMetThrLeuAlaLysGlnGluAlaLysLeuValHisGluLysIleLys 180
DB 525 GAAGAGCGCAACATGACATTTGGCAAGAAAGAGCAAAATACTACACAGCAAGAAATCTCT 584
QY 181 AspLysHisThrAsnAspGluAspValIleArgIleLeuSerThrArgSerLysAlaGln 200
DB 585 GACAAAGGCTTACATATGATGAGAGCTCATCCGAATTTATCTCTAGAGATTAAGCAG 644
QY 201 IleAsnAlaThrPheAsnArgTyrGlnAspAspHisGlyGluGluIleLeuLysSerLeu 220
DB 645 CTGAATGCAACATTTCAACACATCTGCTGACCAACATGGCAGTGAATCAACAAAGCATCTG 704
QY 221 GluGluGluValAspAspAspValPheLeuAlaLeuLeuArgSerThrIleGlnCysLeu 240
DB 705 GAAACT---GATTCGTATGATGATGATCTGAAATTTATCTAGCCACAGCAATAGATGCTTG 761
QY 241 ThrArgProGluLeuValPheValAspValLeuArgSerAlaIleAsnLysThrGlyThr 260
DB 762 AAAAACCAGAGAAACCTTTGAGAAAGTTCTTGATGGCTTACAAAGGTATACAGGACACA 821
QY 261 AspGluGluValaLeuThrArgIleValThrThrArgAlaGluIleAspLeuValIle 280
DB 822 GACGAATGGGACCTTACTTACAGTTGTCTACACTCTGGGCTGAAGTTGACATGCAAGCATATC 881

```

```

QY 281 GlyGluGluTyrGlnArgArgAsnSerIleProLeuGluLysAlaIleThrLysAspThr 300
DB 882 AAAGAAGAGTACCATTAAGAGAAACAGTGTCCATTGACCGTCCAAATTTGCTGAGACACT 941

```

```

QY 301 ArgGlyAspTyrGlnLysMetLeuValAlaLeuLeuGlyGluAspAspAla 317
DB 942 TCAGGAGACTTATGAAAGATGCTTTCGCTTTGATTTGGGCGATGAGATGCT 992

```

# RESULT 13

ADN74790  
 ID ADN74790 standard; cDNA; 954 BP.

XX  
 AC ADN74790;

XX  
 DT 15-JUL-2004 (first entry)

XX  
 DE Thale cress cDNA repressed in E2Fa/Dpa expressing plants SeqID 2655.

XX  
 KW gene; ss; plant; transgenic; E2Fa/Dpa transcription factor;

XX  
 KW growth regulator; animal feed product; thale cress;

XX  
 KW cell wall biosynthesis; nitrogen metabolism; carbon metabolism.

XX  
 PA Arabidopsis thaliana.

XX  
 PN W02004035798-A2.

XX  
 PD 23-APR-2004.

XX  
 PF 20-OCT-2003; 2003WO-BP011658.

XX  
 PR 18-OCT-2002; 2002EP-00079408.

XX  
 PA (CROP-) CROPDESIGN NV.

XX  
 PI Inze D, De Veylder L, Vlieghe K;

XX  
 DR WPI; 2004-348466/32.

XX  
 DR P-PADB; ADN74791.

XX  
 PT Altering plant characteristics, useful for producing plants for enzyme or

XX  
 PT pharmaceutical production comprises modifying in a plant, expression of

XX  
 PT one or more nucleic acids and/or modifying level or activity of one or

XX  
 PT more proteins.

XX  
 PS Claim 1; SEQ ID NO 2685; 134bp; English.

XX  
 CC This invention relates to a novel method for altering one or more plant

XX  
 CC characteristics. Specifically, it refers to identifying genes that are up

XX  
 CC - or down-regulated in transgenic plants overexpressing the heterodimeric

XX  
 CC E2Fa/Dpa transcription factor of Arabidopsis and using these sequences to

XX  
 CC alter plant characteristics accordingly. The present invention describes

XX  
 CC generating transgenic plants for the production of growth regulators,

XX  
 CC enzymes, therapeutics, pharmaceuticals and animal feed products, where

XX  
 CC the altered plant characteristics are selected from increased yield or

XX  
 CC biomass, enhanced survival capacity, stress tolerance, plant architecture

XX  
 CC or physiology, altered endoreplication, biochemistry, signal

XX  
 CC transduction, storage lipid mobilisation and/or altered photosynthesis,

XX  
 CC each relative to the corresponding wild type plants. Accordingly, these

XX  
 CC sequences can also be useful as positive or negative selectable markers

XX  
 CC during transformation of cells or tissues. The identified genes play a

XX  
 CC role in a variety of biological processes such as DNA replication, cell

XX  
 CC wall biosynthesis, nitrogen and/or carbon metabolism or they function as

XX  
 CC transcription factors. This polynucleotide sequence is thale cress cDNA

XX  
 CC repressed 1.3 fold or more in plants overexpressing the E2Fa/Dpa

XX  
 CC transcription factor, given in an exemplification of the invention.

XX  
 SO Sequence 954 BP; 301 A; 216 C; 216 G; 221 T; 0 U; 0 Other;

# Alignment Scores:

Pred. No.:	2.38e-98	Length:	954
Score:	1067.00 <td>Matches:</td> <td>203</td>	Matches:	203
Percent Similarity:	79.81% <td>Conservative:</td> <td>50</td>	Conservative:	50



Best Local Similarity: 64.04% Mismatches: 64  
Query Match: 66.23% Indels: 0  
DB: 12 Gaps: 0  
US-10-690-564-2 (1-317) x ADNT4790 (1-954)

```
QY      1 MetAlaThrLeuLysValSerAspSerValProAlaProSerAspAspAlaGluGlnLeu 20
        |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
DB      1 ATGGCGCTCTCAAGTCCCAAGCAATGTTCTCTTCCGAAATATACGCCGACACTC 60
QY      21 ArgThrAlaPheGluGlyTrpGlyThrAsnGluAspLeuLeileSerileuAlaHis 40
        |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
DB      61 CACAGGCTTTTCAGGATGGGTACCAAGAGAGTGTATCAATATCAATAGCTCAC 120
QY      41 ArgSerAlaGluGlnArgLysValIleArgGlnAlaLeuThrArgLysGluAsp 60
        |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
DB      121 AGAAGCCGACACACGACGCTGATCCGACGCTTATGACGCTACCTACATGAGAT 180
QY      61 LeuLeuLysThrLeuAspLysGluLeuSerAsnAspPheGluArgAlaIleLeuLeuTrp 80
        |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
DB      181 CTCTCAAGCATTTAGCAAGAGCTTTCTAGGACCTTGAGAGACTGTGATGTGG 240
QY      81 ThrLeuGluProGlyGluArgAspAlaLeuLeuAlaAsnGluAlaThrLysArgTrpThr 100
        |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
DB      241 ACTCTGATCCACAGAGAGATGCTTATTTGGCTTAAGAAATCCACCAAGATGTTACCC 300
QY      101 SerSerAsnGlnAlaLeuMetGluValAlaCysThrArgThrSerThrGlnLeuLeuHis 120
        |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
DB      301 AAGAACATTTGGGTCTGTTGAATCGCTTGACAAAGGCTCTCTTGACTTATCAAG 360
QY      121 AlaArgGlnAlaLeuThrAlaArgLysLysSerLeuGluGlnAspValAlaHisHis 140
        |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
DB      361 GTCAGCAAGCTTACCAAGCTTCATACAAAGAAATCAATGAGAGAAATGTCGCAAC 420
QY      141 ThrThrGlyAspPheArgLysLeuLeuValSerLeuValThrSerTrpArgTrpGlyGly 160
        |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
DB      421 ACATCTGGTGACCTTGTAAGCTTGTGCTTCTCTTGTGACCTTTCAGTATGAAGA 480
QY      161 AspGluValaAsnMetThrLeuAlaLysGlnGluAlaLysLeuValHisGluLysIleLys 180
        |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
DB      481 GATGATGTGAACATGATGCTTGCAAGATCTGAGATCACTGACGAGAGGCTCTCA 540
QY      181 AspLysHisTrpAsnAspGluAspValIleArgIleLeuSerThrArgSerLysAlaGln 200
        |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
DB      541 GAGAAATCTTACAGTGCATGATCTTCATCAGATCTTGACAAACAAAGAACACAGC 600
QY      201 IleAsnAlaThrPheAsnArgTrpGlnAspAspHisGlyGluGluIleLeuLysSerLeu 220
        |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
DB      601 CTGGTGCAACACTCTCAACCACTCAACAGAGTATGAAACGCCCATTAACAGAACTTG 660
QY      221 GluGluGlyAspAspAspAspLysPheLeuAlaLeuLeuArgSerThrIleGlnCysLeu 240
        |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
DB      661 AAGGAAGATCGGACGACATGACTACATGAATCTAAGAGCTGTAATCACATGTGG 720
QY      241 ThrArgProGluLeuTrpPheValaAspValaLeuArgSerAlaIleAsnLysThrGlyThr 260
        |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
DB      721 ACATACCTCGAAGACATTTTGAAAGGTTCTTCGCTCATCATCAACAAATGGGACAC 780
QY      261 AspGluGlyAlaLeuThrArgIleValaThrThrArgAlaGluIleAspLeuLysValIle 280
        |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
DB      781 GACGAATGGGAGACTTAACCCGATCGTACACGAACTGAAGTGAATGGAACCATC 840
QY      281 GlyGluGluTrpGlnArgArgAsnSerIleProLeuGluLysValaIleThrLysAspThr 300
        |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
DB      841 AAAGAGGATATATACGAGAAACAGCATTCCTTTGACCGTGTCTATTCGCAAGACACT 900
QY      301 ArgGlyAspTrpGlyLysMetLeuValaIleLeuLeuGlyGluLysAspAla 317
        |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
DB      901 TCTGGTGACTATGAGGACATGCTTGTGCTCTCTGAGACATGGCATGCT 951
```

RESULT 14  
AAC49728  
ID AAC49728 standard; DNA; 1156 BP.

```
XX      XX AAC49728;
AC      AC 18-OCT-2000 (first entry)
XX      XX
DT      DT
DE      DE Arabidopsis thaliana DNA fragment SEQ ID NO: 62219.
XX      XX Hybridization assay; genetic mapping; gene expression control;
KW      KW protein identification; signal transduction pathway; metabolic pathway;
XX      XX promoter; termination sequence; ss.
OS      OS Arabidopsis thaliana.
XX      XX
PN      PN EP1033405-A2.
PD      PD 06-SEP-2000.
XX      XX
XX      XX 25-FEB-2000; 2000EP-00301439.
PF      PF
XX      XX 25-FEB-1999; 99US-0121825P.
PR      PR 05-MAR-1999; 99US-0123180P.
XX      XX 09-MAR-1999; 99US-0122548P.
PR      PR 23-MAR-1999; 99US-0125788P.
XX      XX 25-MAR-1999; 99US-0126264P.
PR      PR 29-MAR-1999; 99US-0126785P.
XX      XX 01-APR-1999; 99US-0127462P.
PR      PR 06-APR-1999; 99US-0128234P.
XX      XX 08-APR-1999; 99US-0128714P.
PD      PD 16-APR-1999; 99US-0128845P.
XX      XX 19-APR-1999; 99US-0130077P.
PR      PR 21-APR-1999; 99US-0130449P.
XX      XX 23-APR-1999; 99US-0130510P.
PR      PR 23-APR-1999; 99US-0130891P.
XX      XX 28-APR-1999; 99US-0131449P.
PR      PR 30-APR-1999; 99US-0132048P.
XX      XX 30-APR-1999; 99US-0132407P.
PR      PR 04-MAY-1999; 99US-0132484P.
XX      XX 05-MAY-1999; 99US-0132485P.
PR      PR 06-MAY-1999; 99US-0132486P.
XX      XX 06-MAY-1999; 99US-0132487P.
PR      PR 07-MAY-1999; 99US-0132863P.
XX      XX 11-MAY-1999; 99US-0134256P.
PR      PR 14-MAY-1999; 99US-0134218P.
XX      XX 14-MAY-1999; 99US-0134219P.
PR      PR 14-MAY-1999; 99US-0134221P.
XX      XX 14-MAY-1999; 99US-0134370P.
PR      PR 18-MAY-1999; 99US-0134768P.
XX      XX 19-MAY-1999; 99US-0134941P.
PR      PR 20-MAY-1999; 99US-0135124P.
XX      XX 21-MAY-1999; 99US-0135353P.
PR      PR 24-MAY-1999; 99US-0135629P.
XX      XX 25-MAY-1999; 99US-0136021P.
PR      PR 27-MAY-1999; 99US-0136392P.
XX      XX 28-MAY-1999; 99US-0136782P.
PR      PR 01-JUN-1999; 99US-0137222P.
XX      XX 03-JUN-1999; 99US-0137528P.
PR      PR 04-JUN-1999; 99US-0137502P.
XX      XX 07-JUN-1999; 99US-0137724P.
PR      PR 08-JUN-1999; 99US-0138094P.
XX      XX 10-JUN-1999; 99US-0138540P.
PR      PR 10-JUN-1999; 99US-0138847P.
XX      XX 14-JUN-1999; 99US-0139119P.
PR      PR 16-JUN-1999; 99US-0139452P.
XX      XX 16-JUN-1999; 99US-0139453P.
PR      PR 17-JUN-1999; 99US-0139492P.
XX      XX 18-JUN-1999; 99US-0139454P.
PR      PR 18-JUN-1999; 99US-0139455P.
XX      XX 18-JUN-1999; 99US-0139456P.
PR      PR 18-JUN-1999; 99US-0139457P.
XX      XX 18-JUN-1999; 99US-0139458P.
PR      PR 18-JUN-1999; 99US-0139459P.
XX      XX 18-JUN-1999; 99US-0139460P.
PR      PR 18-JUN-1999; 99US-0139461P.
```

```
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140685P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144684P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 28-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147319P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.

PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161044P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161820P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Alignment Scores:
Pred. No.: 3,05e-98
Score: 1067,00
Length: 1156
Percent Similarity: 79,81%
Matches: 203
Best Local Similarity: 64,04%
Conservative: 60
Query Match: 66,23%
Mismatches: 54
DB: 3
Indels: 0
Gaps: 0

US-10-690-564-2 (1-317) x AAC49728 (1-1156)
QY 1 MetAlaThrLeuValSerAspSerValProAlaProSerAspAspAlaGluGlnLeu 20
Db 95 ATGGCTCTCTCAAAATCCCAAGCAATGTTCTCTCCCGAAGATACCCGACACTC 154
QY 21 ArgThrAlaPheGluGlyTyrGlyThrAsnGluAspLeuIleIleSerIleLeuAlaHis 40
Db 155 CACAAGCTTTTTCAGATGGGGTACCAAGAAAGCTGATCATCTACTAGCTTAC 214
QY 41 ArgSerAlaGluGlnArgValIleArgGlnAlaTyrHisGluThrTyrGlyGluAsp 60
```



PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0150888P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 03-AUG-1999; 99US-0146389P.  
PR 04-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 06-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.

PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161922P.  
PR 28-OCT-1999; 99US-0161993P.  
PR 29-OCT-1999; 99US-0162142P.

## Alignment Scores:

Pred. No. 3.07e-98 Length: 1160  
Score: 1067.00 Matches: 203  
Percent Similarity: 79.81% Conservative: 50  
Best Local Similarity: 64.04% Mismatches: 64  
Query Match: 66.23% Indels: 0  
DB: 3 Gaps: 0

US-10-690-564-2 (1-317) x AAC40038 (1-1160)

QY 1 MetAlaThrLeuIysValSerAspSerValProAlaProSerAspAlaGluGlnLeu 20  
|||||  
DB 96 ATGGCGCTCTCAAAAGCTTCCCAAGCATGTTCTTCCCGAAGATACCCGCAACTC 155  
|||||  
QY 21 ArgThrAlaPheGluGlyTrpGlyThrAsnGluAspLeuIleIleSerIleLeuAlaHis 40  
|||||  
DB 156 CACAAAGCTTTTTCAGAGTGGGTACCAAGAACTGATCTATCAATCACTACAGCTCAC 215  
|||||  
QY 41 ArgSerAlaGluGlnIArgLysValIleArgGlnAlaTyrHisGluThrTyrGlyGluAsp 60  
|||||  
DB 216 AGGAACGACAGCAACGACGCTGATCCGACGGTTATGACGACTACATGAGAT 275  
|||||  
QY 61 LeuLeuIysThrLeuAspLysGluLeuSerAsnAspPheGluArgAlaIleLeuLeuTrp 80  
|||||  
DB 276 CTCTCAAGCATTAACAAGAGCTTTGTAGCGACTTGAAGAGCTGTGATGTGTGG 335  
|||||  
QY 81 ThrLeuGluProGlyGluArgAspAlaLeuLeuAlaAsnGluAlaThrIysArgTrpThr 100  
|||||  
DB 336 ACTCTTGATCCACCAAGAGAGATGCTTATTGGCTTAAGAAATCCACCAAGATGTTACAC 395  
|||||  
QY 101 SerSerAsnGlnValLeuMetGluValAlaIaCysThrArgThrSerThrGlnLeuLeuHis 120  
|||||  
DB 396 AAGAACAAATGGGTTCTTGTGAATCGCTTCACAAAGCGCTGCTTACGCTTATCAAG 455

```
Qy 121 AlArgGlnAlaIleuThrHisAlaArgTyrLysSerLeuGlnuAspValAlaHisHis 140
Db 456 GTCAGAGCAAGCTTACCAGCTCCATACAGAAATCAATCAGAGAGATGCGCCACAC 515
Qy 141 ThrThrGlyAspPheArgLysLeuValSerLeuValThrSerTyrArgTyrGlnGly 160
Db 516 ACATCTGGTACCTTGTAGCTCTTGTCTCTCTGTGTGACACTTCAGGTATGAAGGA 575
Qy 161 AspGluValAsnMetThrLeuAlaLysGlnGluAlaLysLeuValHisGluLysIleLys 180
Db 576 GATGATGTGAACATGATGCTGCAAGCTGAAGCTAAGTACTTCCAGGAAGTCTCA 635
Qy 181 AspLysHisTyrAsnAspGluAspValIleArgIleLeuSerThrArgSerLysAlaGln 200
Db 636 GAGAAATCCTACAGTGCATGACTTCAATCAGAAATCTTGACACAGAAAGCAAGCAG 695
Qy 201 IleAsnAlaThrPheAsnArgTyrGlnAspAspHisGlyGlnGluIleLeuLysSerLeu 220
Db 696 CTCGTTGCACACACTCAACCACTACACACAGAGATGAAACGCCATTACAAAGAACTTG 755
Qy 221 GluGluGlyAspAspAspAspLysPheLeuAlaLeuLeuArgSerThrIleGlnCysLeu 240
Db 756 AAGGAAAGTCGAGCGAATGACTACATGAAACTACTAAGCTGTAATCACAATGTTG 815
Qy 241 ThrArgProGluLeuTyrPheValAspValIleuArgSerAlaIleAsnLysThrGlyThr 260
Db 816 ACATACCTTGAAGAGCATTTTGAGAGAGTCTTGTCTATCAATCAACAAATGGAGCA 875
Qy 261 AspGluGlyAlaLeuThrArgIleValThrThrArgAlaGluIleAspLeuLysValIle 280
Db 876 GACGAATGGGGACTTAACCGAGTCGTGACTACAGAACTGAAGTTGACATGGAACGATC 935
Qy 281 GlyGluGluTyrGlnArgArgAsnSerIleProLeuGluLysAlaIleThrLysAspThr 300
Db 936 AAAGAGGAATATCAGCGAAGAAACAGCATTCCTTGAGCCGTCTATCGCCAAAGACACT 995
Qy 301 ArgGlyAspTyrGluLysMetLeuValAlaLeuLeuGlyGluAspAspAla 317
Db 996 TCTGGTGACTATGAGCATGCTTGTGCTCTTCTCGGACATGGCATGCT 1046
```

Search completed: August 22, 2005, 10:30:32  
Job time : 677 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 22, 2005, 07:26:26 ; Search time 3736 Seconds

(without alignments)  
3229.757 Million cell updates/sec

Title: US-10-690-564-2

Perfect score: 1611

Sequence: 1 MATLKXSDSPARSDDAEQL.....KOTRGDYEKMLVALLGEDDA 317

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=xlh  
-Q/cgnt2\_1/USPTO.spool/US10690564/runat 19082005 172307 25304/app query.fasta\_1.455  
-DB=EST -QPMF=fastap -SUFFIX=rc -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blomsum62 -TRANS=human40.cdi -LIST=45  
-DOCALLIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFM=pco -NOR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US10690564\_QCGN\_1.1\_3437@runat 19082005 172307 25304 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -MARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6  
-Fgapext=7 -YGAPOP=10 -YGAPEXT=0.5 -DELop=6 -DELext=7

Database :

EST+  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_hic:\*  
4: gb\_est3:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_gsa1:\*  
9: gb\_gsa2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1611	100.0	1095	3	CNSOABR2
2	1579	98.0	1137	3	CNSOABR2
3	1561	96.9	1155	3	CNSOABR2
4	1520	94.4	1034	3	CNSOABR2
5	1502	93.2	910	7	CF652924
6	1489	92.4	1063	5	CF652924
7	1406	87.3	843	5	CF652048
8	1329	82.5	814	6	CA782049
9	1327	82.4	966	3	CNSOABR3

10	1257.5	78.1	782	5	BU635010
11	1228.5	76.3	775	5	BU636350
12	1188.5	73.8	929	6	CB350421
13	1137.5	70.6	889	7	CO121757
14	1111.5	69.0	896	7	CO129429
15	1109.5	68.9	881	7	CO116938
16	1103.5	68.5	855	7	CO116932
17	1095	68.0	860	6	CB256182
18	1084.5	67.3	861	7	CO128796
19	1084.5	67.3	863	7	CO128013
20	1083.5	67.3	870	7	CO132053
21	1082.5	67.2	880	7	CO116065
22	1080.5	67.1	868	7	CO132772
23	1078.5	66.9	864	7	CO128843
24	1077.5	66.9	866	7	CO124740
25	1075	66.7	672	5	BP561140
26	1067.5	66.3	933	7	CK266662
27	1067	66.2	1107	3	CNSOABR1
28	1064.5	66.1	867	6	CB893541
29	1056.5	65.6	979	7	CK277220
30	1048.5	65.1	841	7	CO113247
31	1048.5	65.1	849	7	CO121318
32	1046.5	65.0	843	7	CO121288
33	1046.5	65.0	850	4	BM359402
34	1043.5	64.8	847	6	CB894292
35	1043.5	64.8	901	7	CO128478
36	1039.5	64.5	841	7	CO073880
37	1037	64.4	865	3	CNSOAC11
38	1036.5	64.3	880	7	CO117199
39	1029.5	63.9	903	7	CO074048
40	1027.5	63.8	938	7	CV271171
41	1027	63.7	661	5	BP560951
42	1020.5	63.3	818	7	CO131729
43	1015	62.9	878	7	CNSOABR2
44	1013.5	62.9	915	7	CK263099
45	1013.5	62.9	915	7	CK267919

#### ALIGNMENTS

RESULT 1  
CNSOABR2 1095 bp mRNA linear HTC 06-FPB-2004  
Arabidopsis thaliana Full-length cDNA Complete sequence from clone  
GSLTPGH56ZG07 of Hormone Treated Callus of strain col-0 of  
Arabidopsis thaliana (thale cress).

ACCESSION BX816579.1 GI:42471684  
VERSION BX816579  
KEYWORDS  
SOURCE HTC; GSI CDNA  
ORGANISM Arabidopsis thaliana (thale cress)

REFERENCE  
AUTHORS Menard,M., Craud,C., Quetier,F., Scarpelli,C., Schachter,V.,  
Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.  
TITLE Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences:  
A Combined Approach to Evaluate and Improve Arabidopsis Genome  
Annotation  
Unpublished  
2 (bases 1 to 1095)

JOURNAL  
AUTHORS  
TITLE  
JOURNAL  
COMMENT  
Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :  
Direct Submission  
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr  
- Web : www.genoscope.cns.fr)  
The sequences are based on single pass reads.  
Life Technologies (a division of Invitrogen) members carried out  
full-length libraries construction : Temple G.  
Genoscope members carried out sequencing and annotation : Castelli  
V., Aury J.M., Jallion O., Wincker P., Menard M., Craud C.,

Schachter V., Weissenbach J., Salanoubat M.  
 URGV INRA : Clepet C., Caboche M.  
 Annotation is based on the June 2003 version of the Arabidopsis  
 genome released by MIPS (Munich Information Center for Protein  
 Sequences). 5 prime and 3 prime are assembled with Phrap.  
[http://www.genoscope.cns.fr/externe/sequences/Banque\\_Projet\\_EF/Full](http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full)  
 length  
<http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis>.

## FEATURES

## source

Location/Qualifiers  
 1..1095  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /strain="Col-0"  
 /db\_xref="taxon:3702"  
 /clone="GSLTRGH56ZG07"  
 /tissue\_type="Hormone Treated Callus"  
 /plasmid="pCMVSPORT\_6"  
 1..1095  
 /gene="At1g35720"

## gene

## ORIGIN

## Alignment Scores:

Pred. No.: 2e-180 Length: 1095  
 Score: 1611.00 Matches: 317  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 3 Gaps: 0

US-10-690-564-2 (1-317) x CNSOABR2 (1-1095)

QY 1 MetAlaThrLeuValSerAspSerValProAlaProSerAspAlaGluGluLeu 20  
 Db 23 ATGGCACTTTAAAGTTTCTGATTCCTGCTCTTCTATGATGCTGAGCAATTG 82  
 QY 21 ArgThrAlaPheGluGlyTyrGlyThrAsnGluAspLeuIleSerIleLeuAlaHis 40  
 Db 83 AGAACCGCTTTTGAAGATGGGGTACGACGAGCACTTGTATCAATCAATCTTGCTGAC 142  
 QY 41 ArgSerAlaGluGluAlaGlyValIleArgGluAlaTyrHisGluThrTyrGlyGluAsp 60  
 Db 143 AGAAGTGTCTTAACAGAGAAAGTCATCAGCAAGCAATCACCAAACTACCGCGAAGAC 202  
 QY 61 LeuLeuValThrLeuAspLeuValGluLeuSerAsnAspPheGluAlaIleLeuLeuTyr 80  
 Db 203 CTTCTCAAGACTCTTGAACAGAGACTCTTACGATTTCCAGAGAGCTATCTTGTTGG 262  
 QY 81 ThrLeuGluProGlyGluArgAspAlaLeuLeuAlaAsnGluAlaThrIleValArgTyrPhe 100  
 Db 263 ACTCTGAAACCGGTGAGCGTGAATGCTTTATGCTAATGACACTCAAAAAGATGAGACT 322  
 QY 101 SerSerAsnGluValLeuMetGluValAlaCysThrArgThrSerThrGluLeuLeuHis 120  
 Db 323 TCAGCAACCAAGTCTTATGAGAGTCTTGACAGCAATCAACGACGCTGCTTAC 382  
 QY 121 AlaArgGluAlaTyrHisAlaArgTyrIleValSerLeuGluGluAspValAlaHisHis 140  
 Db 383 GCTAGGCAACCTTACCATGCTCTCCCTACAGAAAGTCTTAAAGAGACGTTGCTTACAC 442  
 QY 141 ThrThrGlyAspPheArgIleLeuLeuValSerLeuValThrSerTyrArgTyrGlyGly 160  
 Db 443 ACTACCGGTGACTTCAGAAAGCTTTGTTGTTCTCTTGTACTCATPACAGTACGAAGA 502  
 QY 161 AspGluValAsnMetThrLeuAlaValysGluGluAlaValysLeuValHisGluValIleHis 180  
 Db 503 GATGAAGTGAACATGACATTTGGCTTAAAGCAAGAGCTTAAGTGTCCATGAGAAATCAAG 562  
 QY 181 AspIleHisIleTyrAsnAspGluAspValIleArgIleLeuSerThrArgSerIleValGlu 200  
 Db 563 GACAGACACTACATGATGAGATGTTATTAAGATCTTGTCCACAGAGCAAAAGCTCAG 622  
 QY 201 IleAsnAlaThrPheAsnArgTyrGluAspAspHisGlyGluGluIleLeuValSerLeu 220

Db 623 ATCAATGCTACTTTTAACCGTTACCAAGATGATCAGCGAGAAATTCCTCAAGACTCTT 682  
 QY 221 GluGluGlyValAspAspAspIlePheLeuAlaLeuLeuValArgSerThrIleGluGlyLeu 240  
 Db 683 GAGGAAGGAGAGATGATGACAAAGTCTCTGACATTTGAGGTCAACCAATTCAGGCTTG 742  
 QY 241 ThrArgProGluLeuTyrPheValAspValIleLeuArgSerAlaIleAsnIleThrIleVal 260  
 Db 743 ACAAGACCAAGACTTACTTCTTGTTCATGTCCTTCTGACGATCAACAAACTGGAACT 802  
 QY 261 AspGluGlyAlaLeuThrArgIleValThrThrArgAlaGluIleAspLeuValIle 280  
 Db 803 GATGAAGGAGCACTCACTGATTTGTGACCAAGAGCTGAGATTGACTTGAAGCTCAATT 862  
 QY 281 GlyGluGlyTyrGluArgArgAsnSerIleProLeuGluValValIleThrIleValPhe 300  
 Db 863 GGAAGAGATACCAAGCGCAGGAACGACATCTTTGGAGAAAGCTATTACCAAGACACT 922  
 QY 301 ArgGlyAspTyrGlyLeuMetLeuValAlaLeuLeuGlyGluAspAspAla 317  
 Db 923 CGTGAAGATTACGAGAAAGATGCTCGTCCGACTTCCGTGAAGATGATGCT 973

RESULT 2  
 CNSOABR2 1137 bp mRNA linear HTC 06-FEB-2004  
 LOCUS Arabidopsis thaliana Full-length cDNA Complete sequence from clone  
 DEFINITION GSLTRGH18ZF02 of Hormone Treated Callus of strain col-0 of  
 Arabidopsis thaliana (thale cress).

ACCESSION BX815936.1 GI:42472184  
 VERSION  
 KEYWORDS  
 SOURCE Arabidopsis thaliana (thale cress)  
 ORGANISM Arabidopsis thaliana

REFERENCE  
 AUTHORS Castelli V., Aury J.M., Jallion O., Wincker P., Clepet C.,  
 Menard M., Cruaud C., Quetier F., Scarpelli C., Schachter V.,  
 Temple G., Caboche M., Weissenbach J. and Salanoubat M.  
 TITLE Whole Genome Sequence Comparisons and Full-Length cDNA Sequences:  
 A Combined Approach to Evaluate and Improve Arabidopsis Genome  
 Annotation

JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 1137)  
 AUTHORS  
 TITLE Direct Submission  
 JOURNAL Submitted (18-NOV-2003) Genoscope - Centre National de Séquençage :  
 BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr  
 - Web : www.genoscope.cns.fr)

COMMENT  
 The sequences are based on single pass reads.  
 Life Technologies (a division of Invitrogen) members carried out  
 full-length libraries construction : Temple G.  
 Genoscope members carried out sequencing and annotation : Castelli  
 V., Aury J.M., Jallion O., Wincker P., Menard M., Cruaud C.,  
 Schachter V., Weissenbach J., Salanoubat M.  
 URGV INRA : Clepet C., Caboche M.  
 Annotation is based on the June 2003 version of the Arabidopsis  
 genome released by MIPS (Munich Information Center for Protein  
 Sequences). 5 prime and 3 prime are assembled with Phrap.  
[http://www.genoscope.cns.fr/externe/sequences/Banque\\_Projet\\_EF/Full](http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full)  
 length  
<http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis>.

## FEATURES

## source

Location/Qualifiers  
 1..1137  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /strain="Col-0"  
 /db\_xref="taxon:3702"  
 /clone="GSLTRGH18ZF02"  
 /tissue\_type="Hormone Treated Callus"  
 /plasmid="pCMVSPORT\_6"  
 1..1137  
 gene

ORIGIN /gene="At1g35720"

Alignment Scores:

Pred. No.:	1,34e-176	Length:	1137
Score:	1579.00	Matches:	313
Percent Similarity:	96.62%	Conservative:	1
Best Local Similarity:	96.31%	Mismatches:	1
Query Match:	98.01%	Indels:	10
DB:	3	Gaps:	1

US-10-690-564-2 (1-317) x CNS0ABT1 (1-1137)

```

QY 3 ThrlenuysValserasp-----SerValProAla 12
   |||||
DB 4 ACCTAAAGATAGAGAAAAATAGAAAGACTCTTACGATTAATAAAATCTGTTCCGCT 63
   |||||
QY 13 ProSerAspAspAlaGluGlnLeuArgThrAlaPheGluGlyTrpGlyThrAsnGluAsp 32
   |||||
DB 64 CTTTCGATGATGCTGAGCAATTCAGAACCGCTTTTGAAGATGGGCTACGAACGAGGAC 123
   |||||
QY 33 LeuIleSerIleLeuAlaHisArgSerAlaGluGlnArgIysValIleArgGlnAla 52
   |||||
DB 124 TTGATCATATCAATCTTGCTCAAGAAAGTGTGAACAGAGAAAGTCATCAGGCAAGCA 183
   |||||
QY 53 TyrHisGluThrTyrGlyIleuAspLeuLeuIleuAspIleuAspIleuSerAspAsp 72
   |||||
DB 184 TACCACCAAACTTACGGGCAAGACCTTCAAGACTTGAACAAGAGCTCTTAAAGAT 243
   |||||
QY 73 PheGluArgAlaIleLeuLeuTrpThrLeuGluProGlyIleuArgAspAlaLeuAla 92
   |||||
DB 244 TTGAGAGAGACTATCTTGTGACCTTGAACCCGCTGAGCTGATGCTTATTTGGCT 303
   |||||
QY 93 AsnGluAlaThrIysArgTrpThrSerSerAsnGlnValLeuMetGluValAlaCysThr 112
   |||||
DB 304 AATGAAGCTACAAAGATGAGACTTCAAGCAACCAAGATCTTATGAAAGTTGCTTGACA 363
   |||||
QY 113 ArgThrSerThrGlnLeuLeuHisAlaArgGlnAlaArgGlyIleuAspIleuSer 132
   |||||
DB 364 AGACATCAACGACGCTGCTTACGCTAGGCAAGCTTACATGCTCGCTCAAGAACTCT 423
   |||||
QY 133 LeuGluIleuAspValAlaHisIleThrThrGlyAspPheArgIleuLeuValSerLeu 152
   |||||
DB 424 CTTGAAGAGACGTTGCTCACACACTACCGGTGACTTCCGAAGCTTTGGTTCTCTT 483
   |||||
QY 153 ValThrSerTyrArgTyrGluGlyAspGluValAsnMetThrLeuAlaIysGlnAla 172
   |||||
DB 484 GTTACCTCATACAGGTACGAAGGAGATGAAGTGAATGACATTTGGCTTAAGAAAGCT 543
   |||||
QY 173 LysLeuValHisGluIleuValIleuAspIleuHisIleThrAsnAspGluAspValIleuArgIle 192
   |||||
DB 544 AACCTGCTCATATGAGAAATCAAGGACACACATCATATATAGATGTTATTAGATC 603
   |||||
QY 193 LeuSerThrArgSerIleuValAlaGlnIleAsnAlaThrPheAsnArgTyrGlnAspAspHis 212
   |||||
DB 604 TTGTCCACAAAGAACCAAGCTCAGATCAATGCTTTTAAACGTTTACCAAGATGATCAT 663
   |||||
QY 213 GlyGluGluIleuLeuIleuSerLeuGluGluGlyAspAspAspIleuPheLeuAlaLeu 232
   |||||
DB 664 GCGAGGAAATTTCTCAAGAGCTTTGAGGAGAGATGATGATGATCAAGATTCCTTGACCTT 723
   |||||
QY 233 LeuArgSerThrIleGlnCysLeuThrArgProGluLeuTrpPheValAspValLeuArg 252
   |||||
DB 724 TTAGGTCAACCACTTCAGTCTTGAACAAGACAGAGCTTTTATCTTGCATGTTCTTCGT 783
   |||||
QY 253 SerAlaIleAsnIysThrGlyThrAspGluGlyValAlaLeuThrArgIleValIleThrArg 272
   |||||
DB 784 TCAGCATCAACAAATTCGAACTGATGAAGGACCACTCACTGAATTTGACACCAAGA 843
   |||||
QY 273 AluGluIleAspLeuIysValIleGlyGluGluIleuTrpGlnArgIleAsnSerIleProLeu 292
   |||||
DB 844 GCTGAGATGATCTTGAAGGTCAATTGAGAGAGATACCAAGCGCAGAAACACATTCCTTTG 903

```

QY 293 GluLysAlaIleThrIysAspThrArgGlyAspTyrGluIleuMetLeuValAlaLeu 312  
 |||||  
 DB 904 GAGAAAGCTATTTCCAAAGACACTCGTGAATTTACGAAAGATGCTCGTCGACTTCTC 963  
 |||||

QY 313 GlyIleAspAspAla 317  
 |||||  
 DB 964 GGTGAAGATGATGCT 978  
 |||||

RESULT 3  
 CNS0AB51  
 LOCUS  
 DEFINITION  
 Arabidopsis thaliana Full-length cDNA Complete sequence from clone GSIR1622B06 of Silique of strain col-0 of Arabidopsis thaliana (thale cress).

ACCESSION 1155 bp mRNA linear HTC 06-FEB-2004  
 VERSION BX818120.1 GI:42471185  
 KEYWORDS HTCC; GSUT cDNA.  
 SOURCE Arabidopsis thaliana (thale cress)  
 ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE  
 1 (bases 1 to 1155)  
 Castelli V., Aury J.M., Jallion O., Wincker P., Clepet C., Menard M., Cruaud C., Queller F., Scarpelli C., Schachter V., Temple G., Caboche M., Weissenbach J. and Salanoubat M., Whole Genome Sequence Comparisons and Full-length cDNA Sequences: A Combined Approach to Evaluate and Improve Arabidopsis Genome Annotation  
 Unpublished  
 2 (bases 1 to 1155)  
 Genoscope.

TITLE  
 JOURNAL  
 REFERENCE  
 AUTHORS  
 TITLES  
 JOURNAL

COMMENT  
 The sequences are based on single pass reads.  
 Life Technologies (a division of Invitrogen) members carried out full-length libraries construction : Temple G.  
 Genoscope members carried out sequencing and annotation : Castelli V., Aury J.M., Jallion O., Wincker P., Menard M., Cruaud C., Schachter V., Weissenbach J., Salanoubat M.  
 URGV INRA : Clepet C., Caboche M.  
 Annotation is based on the June 2003 version of the Arabidopsis genome released by MIPS (Munich Information Center for Protein Sequences) . 5 prime and 3 prime are assembled with Phrap.  
[http://www.genoscope.cns.fr/externe/sequences/Banque\\_Projet\\_EF/FullLength](http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/FullLength)  
<http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis>.

FEATURES  
 source  
 1..1155  
 Location/Qualifiers  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /strain="Col-0"  
 /db\_xref="taxon:3702"  
 /clone="GSIR1622B06"  
 /issue\_type="Silique"  
 /plasmid="pCMVSPORT\_6"  
 1..1155  
 /gene="At1g35720"

ORIGIN  
 Alignment Scores:

Pred. No.:	1.9e-174	Length:	1155
Score:	1561.00	Matches:	310
Percent Similarity:	98.43%	Conservative:	3
Best Local Similarity:	97.48%	Mismatches:	4
Query Match:	96.90%	Indels:	1
DB:	3	Gaps:	0

US-10-690-564-2 (1-317) x CNS0AB51 (1-1155)

QY 1 MetAlaThrLeuLysValSerAspSerValProAlaProSerAspAspAlaGluGlnLeu 20



```

Db      14 ATGGCGACTCTTAAGTTTCTGATTCTGTCGCTCTTCTGATGATGCTGAGCAATTG 73
Qy      21 ArgThrAlaPheGluGluYrTrpGlyThrAsnGluAspLeuIleIleSerIleLeuAlaHis 40
Db      74 AGAACCCGCTTTTGAAGATGGGGTACCAAGAGAGACTGATCATCAATCAATCTTGGCTCAC 133
Qy      41 ArgSerAlaGluGlnArgLysValIleArgGlnAlaTyrHisGluThrTyrGlyGluAsp 60
Db      134 AGAAGTTCTGAACGAGGAAGTCATCAGCCAAAGCATACCAAGAACTTCGCGCGAAGAC 193
Qy      61 LeuLeuYsThrLeuAspLysGluLeuSerAsnAspPheGluArgAlaIleLeuLeuTyr 80
Db      194 CTTCCTCAAGACTCTTGACAAAGAGCTCTTAACGATTTCCAGAGAACTATCTTGTTGG 253
Qy      81 ThrLeuGluProGlyGluArgAspAlaLeuLeuAlaAsnGluAlaThrLysArgTyrPthr 100
Db      254 ACTCTTGAAACCCGGTGGCGTTATGCTTTATGGCTTATATGAACTCAAAAAGATGACT 313
Qy      101 SerSerAsnGlnValLeuMetGluValAlaCysThrArgThrSerThrGlnLeuLeuHis 120
Db      314 TCAGCAACCAAGTCTTATGAAAGTTGCTTGCAACAGACATCAACGACGCTTCCAC 373
Qy      121 AlaArgGlnAlaTyrHisAlaArgTyrLysLysSerLeuGluLysPheValAlaHisHis 140
Db      374 GCTAGGCAAGCTTAACCATGCTCGCTACAGAAAGTTTCTGAAGAGACGTTGCTCACAC 433
Qy      141 ThrThrGluAspPheArgLysLeuLeuValSerLeuValThrSerTyrArgTyrGluGly 160
Db      434 ACTACCGGTACTTCAGAAAGCTTTGGTTTCTTGTTATCCATACAGGTACGAAAGA 493
Qy      161 AspGluValAsnMetThrLeuAlaLysGlnGluAlaLysLeuValHisGluLysIleLys 180
Db      494 GATGAATGAAACATGATGATGGCTTAAGCAAGAAAGCTAGCTGCTCATGAGAAATCAAG 553
Qy      181 AspLysHisTyrAsnAspGluAspValIleArgGluLeuSerThrArgSerLysAlaGln 200
Db      554 GACACAGCACTACATGATGAGATGATTATGAAATCTTGTCACAAAGAACAAAGCTCAG 613
Qy      201 ILAAsnAlaThrPheAsnArgTyrGlnAspAspHisGluGluGluIleLeuLysSerLeu 220
Db      614 ATCAATGCTACTTTTAAACCGTTTACCAAGATGATCAGCGAAGAAATTCACAGAGCTTT 673
Qy      221 GluGluGluLysAspAspAspLysPheLeuAlaLeuLeuArgSerThrIleGlnCysLeu 240
Db      674 GAGGAGAGAGATGATGATGACAACTTCTTGCACTTTGAGTCAACCATTCAGTCTTA 733
Qy      241 ThrArgProGluLeuTyrPheValAspValLeuArgSerAlaIleAsnLysThrGlyThr 260
Db      734 ACAAGACCAAGAGCTTACTTGTGATGATGCTTCGTTCAACAATCAACAACCTGGAAC 793
Qy      261 AspGluGluValAlaLeuThrArgGluIleValThrThrArgAlaGluIleAspLeuLysVal -11 280
Db      794 GATGAAGAGACACTCACTTAAGATTGTGACCAAGAGCTTAATAATTGAAGGTTCCAT 853
Qy      280 eglYgluGluTyrGlnArgArgAsnSerIleProLeuGluLysAlaIleThrLysAspThr 300
Db      854 TGGAGAGAGATACCAAGGACGAGAACGACATTCCTTTGAGAAAGCTTTTCCAAAAAAC 913
Qy      300 rArgGlyAspTyrGluLysMetLeuValAlaLeuLeuGlyGluAspAspAla 317
Db      914 TCGTGAGATTACGAGAAATGCTCTCGCAGCTTCTCGTGAGAGATGATGCT 965

```

```

RESULT 4
CNS0913M      1034 bp      mRNA      linear      HTC 06-FEB-2004
LOCUS      Arabidopsis thaliana Full-length cDNA Complete sequence from clone
DEFINITION      GSLTUS23B01 of Adult vegetative tissue of strain col-0 of
ACCESSION      Arabidopsis thaliana (thale cress).
VERSION      BX841810.1 GI:42454468
KEYWORDS      HTC; GSLT_cDNA.
SOURCE      Arabidopsis thaliana (thale cress)

```

```

ORGANISM      Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosids; eustosids II; Brassicales; Brassicaceae; Arabidopsids.
REFERENCE
AUTHORS      1 (bases 1 to 1034)
Castelli,V., Aury,J.M., Jaillon,O., Wincker,P., Clepet,C.,
Menard,M., Cruaud,C., Quetier,F., Scarpelli,C., Schachter,V.,
Temple,G., Caboche,M., Weissenbach,J. and Salanoubat M.
TITLE      Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences:
A Combined Approach to Evaluate and Improve Arabidopsis Genome
Annotation
JOURNAL      Unpublished
REFERENCE     2 (bases 1 to 1034)
Genoscope.
AUTHORS      Direct Submission
TITLE      Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
JOURNAL      BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr
COMMENT
The sequences are based on single pass reads.
Life Technologies (a division of Invitrogen) members carried out
full-length libraries construction : Temple G.
Genoscope members carried out sequencing and annotation : Castelli
V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,
Schachter V., Weissenbach J., Salanoubat M.
URGV INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis
genome released by MIPS (Munich Information center for Protein
Sequences). 5 prime and 3 prime are assembled with Phrap.
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full
length
http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.
FEATURES
source      1..1034
location/Qualifiers
/mol_type="mRNA"
/strain="Col-0"
/db_xref="taxon:3702"
/clone="GSLTUS23B01"
/tissue_type="Adult vegetative tissue"
/plasmid="pCMVSPORT_6"
1..1034
/gene="Atg35720"
gene
ORIGIN
Alignment Scores:
Pred. No.:      1 21e-169      Length:      1034
Score:          1520.00      Matches:      302
Percent Similarity: 98.41%      Conservative: 7
Best Local Similarity: 96.18%      Mismatches: 5
Query Match:    94.35%      Indels:      1
DB:              3      Gaps:      0
US-10-690-564-2 (1-317) x CNS0913M (1-1034)
Qy      1 MetAlaThrLeuLysValSerAspSerValProAlaProSerAspAspAlaGluGlnLeu 20
Db      23 ATGGGAGACTTAAAGTTTCTGATTCGTGCTGCTTCCTTGAGATGAGCTGAGCAATTG 82
Qy      21 ArgThrAlaPheGluGluYrTrpGlyThrAsnGluAspLeuIleIleSerIleLeuAlaHis 40
Db      83 AGAACCCGCTTTTGAAGATGGGGTACCAAGAGAGCTGATCATCAATCTTGGCTCAC 142
Qy      41 ArgSerAlaGluGlnArgLysValIleArgGlnAlaTyrHisGluThrTyrGlyGluAsp 60
Db      143 AGAAGTCTGAACAGAGAAAGTCATCAGGCAAGCATACCAAGAGACTTACGCGCAAGAC 202
Qy      61 LeuLeuYsThrLeuAspLysGluLeuSerAsnAspPheGluArgAlaIleLeuLeuTyr 80
Db      203 CTTCCTCAAGACTCTTGACAAAGAGGCTCTTAACGATTTGAGAGAGCTATCTTGTTGG 262
Qy      81 ThrLeuGluProGlyGluArgAspAlaLeuLeuAlaAsnGluAlaThrLysArgTyrPthr 100
Db      263 ACTCTTGAAACCCGGTGGCGGTGATGCTTTATGGCTTATGAGAGCTTAAGAAAGATGACT 322

```

Qy 101 SerSerAnGlnValIleuMetGluValAlaCyThrArgThrSerThrGlnIleuLeuHis 120  
 Db 323 TCAAGCAACCAAGTCTTATGGAGGTCTTGCAAGAGACATACAGCGAGCTGCTTAC 382  
 Qy 121 AlaaGlnAlaTyRHISAlaArgTyRlySerLeuGluIleuValAlaHis 140  
 Db 383 GCTAGGCAAGCTTACATGCTCGCTCAAGAGGCTCTTGAGAGGAGCGTTGCTCACAC 442  
 Qy 141 ThrThrGlyAspPheArgLysLeuLeuValSerLeuValThrSerTyArgTyGluGly 160  
 Db 443 ACTACCGGTGACTTCAGAGAGCTTTGGTTCTCTTGTTACTCATACAGTACGAAGA 502  
 Qy 161 AspGluValIleuMetThrLeuAlaLysGlnGluAlaLysLeuValHisGluLysIleLys 180  
 Db 503 GATGAAGTGAACATGACATGGCTTACCAAGAACCTAAGCGGTGCATGAGAGATCAG 562  
 Qy 181 AsplyHisTyRanAspGluAspValIleArgTyleuSerThrArgSerLysAlaGln 200  
 Db 563 GACAGACATACATGATGAGATGTTATTGATCTTGTCCACAGAGAGAGCTCAG 622  
 Qy 201 IleAsnAlaThrPheAsnArgTyRGlAspAspHisGlyGluLysIleuLysSerLeu 220  
 Db 623 ATCAATCTACTTTGGCCGCTTACCAAGATGATCATGCGAGGGAATTCACAGAGCTT 682  
 Qy 221 GluGluGlyAspAspAspAspLysPheLeuAlaLeuLeuArgSerThrIleGlnCysLeu 240  
 Db 683 GAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 742  
 Qy 241 ThrArgProGluLeuTyRPhaValAspValLeuArgSerAlaIleuLysThrGlyThr 260  
 Db 743 ACAAGACCAAGCTTCTTCTTGATGATGATGATGATGATGATGATGATGATGATGATG 802  
 Qy 261 AspGluGlyAlaLeuThrArgIleValThrThrArgAlaGluIleAspLeuLysValIle 280  
 Db 803 GATGAGGAGACACTCACTAGATGTTGTGACCAAGAGCTGATGATGATGATGATGATG 862  
 Qy 281 GlyGluGlyLysPheArgLysValIleProLeuGluLysAlaIleThrLysAspThr 300  
 Db 863 GAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 922  
 Qy 301 ArgGlyAspTyRglLysMetLeuValAlaLeuLeuGlyGlu 314  
 Db 923 CTGAGATTAACAAGAGAG-CTGCTGCACTACTCATGAG 963

RESULT 5  
 LOCUS CF652924 910 bp mRNA linear EST 06-NOV-2003  
 DEFINITION 85-1020135W-066-001-121-SP6P MP12-ADIS-066 Arabidopsis thaliana  
 ACCESSION CDNA clone MP12p20011211Q 5-PRIME, mRNA sequence.  
 VERSION CF652924  
 KEYWORDS GI:37429878  
 SOURCE EST.  
 ORGANISM Arabidopsis thaliana (thale cress)  
 Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 1 (bases 1 to 910)  
 Schmidt,K.J., Soerensen,T.R., Stracke,R., Torjek,O., Altmann,T.,  
 Mitchell-Olds,T. and Weisshaar,B.  
 TITLE large-scale identification and analysis of genome-wide  
 JOURNAL single-nucleotide polymorphisms for mapping in Arabidopsis thaliana  
 MEDLINE Genome Res. 13 (6), 1250-1257 (2003)  
 PUBMED 22683290  
 COMMENT Contact: Weisshaar B  
 ADIS DNA core facility at MP12  
 Max-Planck-Institute for Plant Breeding Research  
 Carl-von-Linne Weg 10, 50829 Koeln, Germany  
 Fax: 00492215062851  
 Email: weisshaar@mpiz-koeln.mpg.de  
 Insert Length: 910 Std Error: 0.00

Plate: 1 row: 1 column: 21  
 Seq primer: SP6P.  
 Location/Qualifiers  
 1. 910  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /ecotype="Ws-0"  
 /db\_xref="GABI:939529"  
 /db\_xref="taxon:3702"  
 /clone="MP12p20011211Q"  
 /isue\_type="root"  
 /lab\_host="E. coli TOP10"  
 /clone\_lib="MP12-ADIS-066"  
 /note="vector: pCMVSPORT6; Site 1: SalI; Site 2: NotI;  
 cDNA library from Arabidopsis thaliana, accession  
 Maesliewakija-0; roots from three weeks old plants grown  
 on MS-plates at 25M-0C with 16 hours light/day; library  
 was made at the Max-Planck-Institute for Plant Breeding  
 Research, Cologne, Germany; cloning sites SalI-NotI,  
 primer sites and orientation:  
 SP6-SalI-CCACGCGTCCG-Sp1ime-cDNA-polyA-CC-NotI-T7; GATEWAY  
 compatible; Note: Sequencing granted in the context of the  
 GABI Arabidopsis Verbund I: Genetic Diversity.  
 'Establishment of high-efficiency SNP-based mapping tools  
 and development of methods for genome-wide mutation  
 detection' PI: Bernd Weisshaar Sequence submission managed  
 by RZPD/GABI-Primary database: <http://gabi.rzpd.de> This  
 clone is available from RZPD; contact RZPD (clone@rzpd.de)  
 for further information."

ORIGIN  
 Alignment Scores:  
 Pred. No.: 1.4e-167 Length: 910  
 Score: 1502.00 Matches: 296  
 Percent Similarity: 99.66% Conservative: 0  
 Best Local Similarity: 99.66% Mismatches: 1  
 Query Match: 93.23% Indels: 0  
 DB: 7 Gaps: 0

US-10-690-564-2 (1-317) x CF652924 (1-910)

Qy 1 MetAlaThrLeuLysValSerAspSerValProAlaProSerAspAspAlaGluGlnIleu 20  
 Db 18 ATGGCACTCTTAAGGTTCTGATTTCTGTTCTGCTCTTGTGATGATCGTGCAATTG 77  
 Qy 21 ArgThrAlaPheGluGlyTyRglTyRanGluAspLeuIleSerIleuAlaHis 40  
 Db 78 AGAACCGCTTTTGAAGAGATGGGTGACGAGCACTTATCATCATCTTGGCTCAC 137  
 Qy 41 ArgSerAlaGluGlnArgLysValIleArgGlnAlaTyRHISGluThrTyRglGluAsp 60  
 Db 138 AGAAGTGTGAAACAGAGAAAGTCAATCAGCAAGCATACATGAACCTACGCGCAAGAC 197  
 Qy 61 LeuLeuLysThrLeuAspLysGluLeuSerAsnAspPheGluArgAlaIleLeuLeuTrp 80  
 Db 198 CTTCACAACTCTTGAACAAGAGCTCTTAAAGACTTGAAGAGACTATCTTGTGTGG 257  
 Qy 81 ThrLeuGluProGlyLysArgAspAlaLeuAlaAsnGluAlaThrLysArgTyRThr 100  
 Db 258 ACTCTTGAACCCGGTGAAGGTATGCTTATTTGGCTTAAGAAAGCTACAAAAGATGACT 317  
 Qy 101 SerSerAnGlnValIleuMetGluValAlaCyThrArgThrSerThrGlnLeuLeuHis 120  
 Db 318 TCAAGCAACCAAGTCTTATGGAGGTCTTGCAAGAGACATACAGCGAGCTGCTTAC 377  
 Qy 121 AlaaGlnAlaTyRHISAlaArgTyRlySerLeuGluIleuValAlaHisHis 140  
 Db 378 GCTAGGCAAGCTTACATGCTCGCTCAAGAGGCTCTTGAGAGGAGCGTTGCTCACAC 437  
 Qy 141 ThrThrGlyAspPheArgLysLeuLeuValSerLeuValThrSerTyArgTyRglGly 160  
 Db 438 ACTACCGGTGACTTCAGAGAGCTTTGGTTCTCTTGTTACTCATACAGTACGAAGA 497

QY 161 AspGluValAsnMetThrLeuAlaLysGlnGluAlaLysLeuValHisGluLysIleLys 180  
| | | | |  
| | | | |  
Db 498 GATGAAGTGAACATGACATTGGCTAAGCAGAAAGCTAAGTGGTCCATGAGAAATCAAG 557  
QY 181 AspLysHisThrAsnAspGluAspValIleArgIleLeuSerThrArgSerLysAlaGln 200  
| | | | |  
| | | | |  
Db 558 GACAGACACTACATCATGTATGAGATGTTATTAAGATCTTGTCCACAAGAACCAAGCTCAG 617  
QY 201 ILAsnAlaThrPheAsnArgTyrGlnAspAspHisGluGluGluIleLeuLysSerLeu 220  
| | | | |  
| | | | |  
Db 618 ATCAATGCTACTTTTACCGTTACCAAGATGATATGCGCAGAAATTTCCAAAGACTTT 677  
QY 221 GlnGluGluLysAspAspAspLysPheLeuAlaLeuLysArgSerThrIleGlnCysLeu 240  
| | | | |  
| | | | |  
Db 678 GAGGAAGAGATGATGATGATGACAAAGTTCCCTTGTGTGAGGTCAACCAATTCAGTGCCTTG 737  
QY 241 ThrArgProGluLeuLysTyrPheValAspValLeuLysSerAlaIleAsnLysThrGlyThr 260  
| | | | |  
| | | | |  
Db 738 ACAAGACCAAGAGCTTACTTTCGATGATCTTCCTTCGTCAGCAATCAACAAACTGGAACCT 797  
QY 261 AspGluGluValaLeuThrArgIleValIleThrThrArgAlaGluIleAspLeuLysValIle 280  
| | | | |  
| | | | |  
Db 798 GATGAAGAGACTCTCACTAATATTGTCACCAAGAGCTGAGATTGACTTGAAGGTCAATT 857  
QY 281 GlyGluGluLysTyrGlnArgAsnSerIleProLeuGluLysAlaIleThr 297  
| | | | |  
| | | | |  
Db 858 GGAAGAGAGTACCAAGCGCAGAAACAGCATTCCTTTGGAGAAAGCCATTACC 908

RESULT 6  
BX838972  
LOCUS  
DEFINITION  
BX838972 Arabidopsis thaliana flowers and buds Col-0 Arabidopsis  
thaliana cDNA clone GSLTFRB502D03 SPRIM, mRNA sequence.  
BX838972  
BX838972.1 GI:42533055

ACCESSION  
VERSION  
SOURCE  
KEYWORDS  
ORGANISM  
EST.  
Arabidopsis thaliana (thale cress)  
Arabidopsis thaliana  
Eukaryotes; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis;  
1 (bases 1 to 1063)  
Castelli,V., Aury,J.M., Jallion,O., Wincker,P., Clepet,C.,  
Menard,M., Cruaud,C., Quetier,F., Scarpelli,C., Schachter,V.,  
Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.  
Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences:  
A Combined Approach to Evaluate and Improve Arabidopsis Genome  
Annotation  
Unpublished (2004)

JOURNAL  
COMMENT  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
The sequences are based on single pass reads.  
Life Technologies (a division of Invitrogen) members carried out  
full-length libraries construction : Temple G.  
Genoscope members carried out sequencing and annotation : Castelli  
V., Aury J.M., Jallion O., Wincker P., Menard M., Cruaud C.,  
Schachter V., Weissenbach J., Salanoubat M.  
JRGV INRA : Clepet C., Caboche M.  
Annotation is based on the June 2003 version of the Arabidopsis  
genome released by MIPS (Munich Information center for Protein  
Sequences).  
http://www.genoscope.cns.fr/externe/sequences/Banque\_Projet\_FR/EST  
http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.

FEATURES  
source  
1.1063  
Location/Qualifiers  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/ecotype="Col-0"  
/db\_xref="taxon:3702"  
/clone="GSLTFRB502D03"  
/tissue\_type="flowers and buds"

ORIGIN  
/clone\_1ib="Arabidopsis thaliana flowers and buds Col-0"  
Alignment Scores:  
Pred. No.: 6,11e-166 Length: 1063  
Score: 1489.00 Matches: 292  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 92.43% Indels: 0  
DB: 5 Gaps: 0  
US-10-690-564-2 (1-317) x BX838972 (1-1063)  
QY 26 G|YTPG|YThrAsnGluAspLeuIleIleSerIleLeuAlaHisArgSerAlaGln 45  
| | | | |  
| | | | |  
Db 2 GGATGGGGTACGACAGAGAGCTTGATCATATCAATCTTGGCTCAGAGAGTCTGACAG 61  
QY 46 ArgLysValIleArgGlnAlaTyrHisGluThrTyrGlyGluAspLeuLysThrLeu 65  
| | | | |  
| | | | |  
Db 62 AGGAAAGTCATCAGCGACATACCAAGAACTACGGCGAAGACCTTCTCAAGACTCTT 121  
QY 66 AspLysGluLeuSerAspAspPheGluArgAlaIleLeuLysThrLeuGluProGly 85  
| | | | |  
| | | | |  
Db 122 GACAGAGAGCTCTTACGATTTCCAGAGAGCTATCTGTGTGAGACTTTGAACCCGAT 181  
QY 86 GluArgAspAlaLeuLeuAlaAsnGluAlaThrLysArgTyrThrSerSerAsnGlnVal 105  
| | | | |  
| | | | |  
Db 182 GACCGTGAAGCTTTATTTGGCTTAATGAGCTCAAAAAGATGAGACTTCAAGCAACCAAGT 241  
QY 106 LeuMetGluValAlaCysThrArgThrSerThrGlnLeuLeuHisAlaArgGlnAlaTyr 125  
| | | | |  
| | | | |  
Db 242 CTTATGGAAGTGGCTTGGCCACAGAGACATCAAGCGAGCTTCACGCTTACGAGACCTTAC 301  
QY 126 HisAlaArgTyrLysSerLeuGluGluAspValAlaHisHisThrThrGlyAspPhe 145  
| | | | |  
| | | | |  
Db 302 CAGTCCGCTACAGAAAGTCTCTTAAGAGAGCGTTCACACACATACCGGTGCTTC 361  
QY 146 ArgLysLeuLeuValSerLeuValThrSerTyrArgTyrGluGlyAspGluValAsnMet 165  
| | | | |  
| | | | |  
Db 362 AGAAAGCTTTTGCTTCTCTTCTTACCTCATACAGTACGAAAGAGATGAAGCAATG 421  
QY 166 ThrLeuAlaLysGlnGluAlaLysLeuValHisGluLysIleLysAspLysHisTyrAsn 185  
| | | | |  
| | | | |  
Db 422 ACATTGGCTTAACCAAGAGCTAAGCTGCTCAGTGAAGAAATCAAGAGACATACAT 481  
QY 186 AspGluAspValIleArgIleLeuSerThrArgSerLysValAlaGlnIleAsnAlaThrPhe 205  
| | | | |  
| | | | |  
Db 482 GATGAGAGATGTTATTAAGATCTTGTCCACAAAGAGCAAGCTCAAGTCAATCTACTTT 541  
QY 206 AsnArgTyrGlnAspAspHisGluGluGluIleLeuLysSerLeuGluGluLysAspAsp 225  
| | | | |  
| | | | |  
Db 542 AACCGTTACCAAGATGATGATGCGAGGAAATTCACAGAGCTTGAGGAAAGAGATGAT 601  
QY 226 AspAspLysPheLeuAlaLeuLysArgSerThrIleGlnCysLeuThrArgProGluLeu 245  
| | | | |  
| | | | |  
Db 602 GATGACAAATTCCTTGACATTTTGGAGTCAACCAATTCAGTGCCTTGAACAGCAAGAGCTT 661  
QY 246 TyrPheValAspValLeuLysSerAlaIleAsnLysThrGlyThrAspGluGluValaLeu 265  
| | | | |  
| | | | |  
Db 662 TACTTGTGATGATCTTCTGTTCCAGATCAACAAAGCTGGAAGCTGATGAAGAGAGACTC 721  
QY 266 ThrArgIleValIleThrArgAlaGluIleAspLeuLysValIleGluGluLysTyrGln 285  
| | | | |  
| | | | |  
Db 722 ACTAAGATTTGTACCAAGAGCTGATGATGACTTGAAGTATATGAGAGAGATACAG 781  
QY 286 ArgArgAsnSerIleProLeuGluLysAlaIleThrLysAspThrArgGlyAspTyrGlu 305  
| | | | |  
| | | | |  
Db 782 CCAGAGAACAGATTCCTTGGAGAAAGCTATTACCAAGACATCGTGAAGATTACGAG 841  
QY 306 LysMetLeuValAlaLeuLeuGluGluAspAspAla 317  
| | | | |  
| | | | |  
Db 842 AAGATGCTGTCGACCTTCGCTGAAGATGATGCT 877

RESULT 7  
CF652048  
LOCUS  
DEFINITION  
30-L020380-066-004-L08-SP6P MP1Z-ADIS-066 Arabidopsis thaliana cDNA  
CF652048  
VERSION  
KEYWORDS  
CF652048.1 GI:37428171  
EST.  
SOURCE  
Arabidopsis thaliana (thale cress)  
ORGANISM  
Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsie.  
REFERENCE  
1 (bases 1 to 843)  
Schmid, K.J., Sørensen, T.R., Stracke, R., Torjek, O., Altmann, T.,  
Mitchell-Olds, T. and Weisshaar, B.  
TITLE  
Large-scale identification and analysis of genome-wide  
single-nucleotide polymorphisms for mapping in Arabidopsis thaliana  
JOURNAL  
Genome Res. 13 (6), 1250-1257 (2003)  
MEDLINE  
22683290  
PUBMED  
12799357  
COMMENT  
Contact: Weisshaar B  
ADIS DNA core facility at MP1Z  
Max-Planck-Institute for Plant Breeding Research  
Carl-von-Linne Weg 10, 50829 Koeln, Germany  
Fax: 00492215062851  
Email: weisshaar@mpiz-koeln.mpg.de  
Insert Length: 843 Std Error: 0.00  
Plate: 4 row: L column: 08  
Seq primer: SP6P.  
Location/Qualifiers  
1..843  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/ecotype="Ws-0"  
/db\_xref="GABI:938637"  
/db\_xref="taxon:3702"  
/clone="MP1Zp2001L084Q"  
/cissue\_type="root"  
/lab\_host="E. coli TOP10"  
/clone\_1db="MP1Z-ADIS-066"  
/note="Vector: PCMVSPORT6; Site 1: SalI; Site 2: NotI;  
cDNA library from Arabidopsis thaliana, accession  
MasJlewskija-0; roots from three weeks old plants grown  
on WS-plants at 26M-0C with 16 hours light/day; library  
was made at the Max-Planck-Institute for Plant Breeding  
Research, Cologne, Germany; cloning sites SalI-NotI,  
primer sites and orientation:  
SP6-Sali-CCACGCGCCG-5prime-cDNA-polyA-CC-NotI-T7; GATEWAY  
compatible; Note: Sequencing granted in the context of the  
GABI Arabidopsis Verbund II: Genetic Diversity,  
'Establishment of high-efficiency SNP-based mapping tools  
and development of methods for genome-wide mutation  
detection'. PI: Bernd Weisshaar Sequence submission managed  
by RZPD/GABI-Primary database: <http://gabi.rzpd.de> This  
clone is available from RZPD; contact RZPD ([clone@rzpd.de](mailto:clone@rzpd.de))  
for further information."

ORIGIN  
Alignment Scores:  
Pred. No.: 3,24e-156 Length: 843  
Score: 1406.00 Matches: 276  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 87.27% Indels: 0  
DB: 7 Gaps: 0  
US-10-690-564-2 (1-317) x CF652048 (1-843)

QY 1 MetAlaThrLeuIyValSerAspSerValProAlaProSerAspAlaGluInLeu 20  
DB 14 ATGGCGACCTTAAGGTTTCTGATTCTGCTCTCTCTGATGATGCTGACCAATTG 73

QY 21 ArgThrAlaPheGluGluTyrProGlyThrAsnGluAspLeuIleIleSerIleLeuAlaHis 40  
DB 74 AGAACCGCTTTTAAAGATGGGGTACCAACGACCTTGATCTATCAATCTTGGCTCAC 133  
QY 41 ArgSerAlaGluGlnArgValIleArgGlnAlaTyrHisGluThrTyrGlyGluAsp 60  
DB 134 AGAAGTGTGAAACAGAGAAAGTCAATCAGGCAAGCATCATTGAACTGCGCGAAGAC 193  
QY 61 LeuLeuIyThrLeuAspIyGluLeuSerAsnAspPheGluAlaIleLeuLeuTyr 80  
DB 194 CTCTCAAGACTCTTACCAAGAGGCTCTCAAGACTCTGAGAGACTATCTTGTGTGG 253  
QY 81 ThrLeuGluProGlyGluArgAspAlaLeuLeuAlaAsnGluAlaThrIyAspGlyProThr 100  
DB 254 ACTCTTGAACCCGTGAGCTGATGCTTTATTTGGTTAAGAACCTTCAAAAATGAGACT 313  
QY 101 SerSerAsnGluValLeuMetGluValAlaCysThrArgThrSerThrGlnLeuLeuHis 120  
DB 314 TCAGCAACCAAGTGTATGGAAGTTGCTTGCAAGAGCATCAACGAGCTGCTTAC 373  
QY 121 AlaArgGlnAlaTyrHisAlaArgTyrIyIySerLeuGluGluAspValAlaHisHis 140  
DB 374 GCTTAGCAAGCTTACCATGCTGCTCAAGAACTCTTGAAGAGACCTTGCTCACAC 433  
QY 141 ThrThrGlyAspPheArgIySerLeuLeuValSerLeuValThrSerTyrArgTyrGly 160  
DB 434 ACTACCGGTGACTCGAAGAGCTTGTGTTCTCTTGTTACCTCAACAGTGAAGGA 493  
QY 161 AspGluValAsnMetThrLeuAlaIyGlnGluAlaIySerValHisGluIyIyIy 180  
DB 494 GATGAAGTGAACATGACATTTGGTTAAGCAAGAGCTTACGTGCTCATGAAATCAAG 553  
QY 181 AspIyHisIyTyrAsnAspGluAspValIleArgIleLeuSerThrArgSerIyAlaGln 200  
DB 554 GACAAAGCTTACCATATGATGATGATGATTTAGATCTTGCCCAAGAAAGCAAGCTCAG 613  
QY 201 IleAsnAlaThrPheAsnArgTyrGlnAspAspHisGlyGluGluIyLeuIySerLeu 220  
DB 614 ATTAATGCTTACTTTAACCGTTTACCAAGATGATCAGGAGAAATTTCTCAAGACTT 673  
QY 221 GluGluGluIyAspAspAspAspIyPheLeuAlaLeuLeuArgSerThrIleGlnCysLeu 240  
DB 674 GAGGAAGAGATGATATATACAAAGTCTCTTGTTTATAGGTCAACCATCTGCTGCTG 733  
QY 241 ThrArgProGluLeuTyrPheValAspValLeuArgSerAlaIleAsnIyThrGlyThr 260  
DB 734 ACAAGACCAAGAGCTTACTTGTGCGATGTTCTTCTGTCAGCAATCAACAAACTGGAAC 793  
QY 261 AspGluGluIyAlaLeuThrArgIyIleValThrThrArgAlaGluIleAsp 276  
DB 794 GATGAAGAGCACTCACTGAATTTGCACCAAGAGCTGAGATTGAC 841

RESULT 8  
CA782049  
LOCUS  
DEFINITION  
01B1B1AR Infected Arabidopsis Leaf Arabidopsis thaliana cDNA, mRNA  
CA782049 814 bp mRNA linear EST 03-DEC-2002  
VERSION  
CA782049.1 GI:26020095  
KEYWORDS  
EST.  
SOURCE  
Arabidopsis thaliana (thale cress)  
ORGANISM  
Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsie.  
REFERENCE  
1 (bases 1 to 814)  
Lundgaard, M., Emmersen, J., Nielsen, K.L., Wilson, I., Somerville, S.  
and Weindler, K.G.  
TITLE  
EST sequencing of Erysiphe cichoracearum infected Arabidopsis  
JOURNAL  
Unpublished (2002)  
COMMENT  
Contact: Karen G. Weindler  
Institut für Biotechnologie

Aalborg Universitet  
 Sohngaardsholmsvej 49, 9000 Aalborg, Denmark  
 Tel: +45 96358467  
 Fax: +45 98141808  
 Email: kym@bio.au.dk

## FEATURES

## source

Location/Qualifiers  
 1. 814  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /ecotype="Columbia"  
 /db\_xref="taxon:3702"  
 /dev\_stage="Plant 3 weeks old, three days post infection"  
 /clone\_lib="Infected Arabidopsis leaf"  
 /note="Organ: Leaf; Vector: pBluescript; Mixed cDNA library of Arabidopsis and E. coli/acetanum infected leaf from three weeks old Arabidopsis plants. Plants were harvested 3 days after infection and mRNA oligo dt selected."

## ORIGIN

## Alignment Scores:

Pred. No.: 4,39e-147 Length: 814  
 Score: 1329.00 Matches: 267  
 Percent Similarity: 98.89% Conservative: 0  
 Best Local Similarity: 98.89% Mismatches: 2  
 Query Match: 82.50% Indels: 2  
 DB: Gaps: 0

US-10-690-564-2 (1-317) x CA782049 (1-814)

QY 1 Meta1aThrLeuValSerAspSerValProAlaProSerAspAspAlaGluGluLeu 20  
 Db 7 ATGGCGACTCTTAG-GTTCTGATTCCTGCTCCTTCGATGATGTCGACATGG 65  
 QY 21 ArgThrAlaPheGluGlyTrpGlyThrAsnGluAspLeuIleIleSerIleLeuAlaHis 40  
 Db 66 AGAACCGCTTTTAAAGAGATGGGTACGACAGACACTGTGATCATCAATCTTGCTCAC 125  
 QY 41 ArgSerAlaGluGlnArgValIleArgGlnAlaTythrIleGluThrTyrglyGluAsp 60  
 Db 126 AGAAGTGTCTAACAAGAGAGATCATCAGGCAAGCATCCACAGAACTTCGGCGAAGAC 185  
 QY 61 LeuLeuValThrLeuAspGluLeuSerAspPheGluAlaIleLeuLeuTrp 80  
 Db 186 CTTCTCAAGACTCTTTCAGACAGAGCTCTCAACGATTCAGAGAGCTATCTTGTTGG 245  
 QY 81 ThrLeuGluProGlyGluArgAspAlaLeuLeuAlaSerGluAlaThrIleValTrpThr 100  
 Db 246 ACTCTTGAACCCGGTGGCGTGAATGCTTAATGGCTAATGAAGCTACGAAAGATGAGCT 305  
 QY 101 SerSerAsnGlnAlaLeuMetGluValAlaCysThrArgThrSerThrglnLeuLeuHis 120  
 Db 306 TCAGGACCAACAGTTCTTAATGAAAGTTGCTTGCACAGACATCAACGACGCTCTTACC 365  
 QY 121 AlaArgGlnAlaTythrIleAlaArgTyrglyLeuSerLeuGluGluAspValAlaHisHis 140  
 Db 366 GCTAGGCAAGCTTAACATGCTCGCTCAAGAAAGCTCTTGAAGAGCGTGTCTCACAC 425  
 QY 141 ThrThrGlyAspPheArgGlyLeuLeuValSerLeuValThrSerTyraArgTyrglyGly 160  
 Db 426 ACTACCGGTACTTCGAAAGCTTTGGTTCTCTGTTTACCCCAACACAGGTACGAAGA 485  
 QY 161 AspGluValAsnMetThrLeuAlaLysGlnGluAlaLysLeuValHisGluValIleLys 180  
 Db 486 GATGAAGTGAACATGACATTTGGCTTAAGCAAGAAAGCTTAAGTGTTCATGAAATCAAG 545  
 QY 181 AspLysHisTyraAsnAspGluAspValIleArgIleLeuSerThraArgSerIleValGln 200  
 Db 546 GACAAAGCATCAATGATGAGATGATTATTAATCTTGTCCACAAAGAACCAAGCTCAG 605  
 QY 201 TLeaAsnAlaThrPheAsnArgTyrglnAspAspHisGlyGluGluIleLeuLysSerLeu 220  
 Db 606 ATCAATGCTACTTTTAAACGTTTACCAAGATGATCAGCGAGAAATTTCTCAAGAGTCTT 665

QY 221 GIUGLUGLYAspAspAspAspPheLeuAlaLeuLeuArgSerThrIleGlnCysLeu 240  
 Db 666 GAGAAAGAGATGATGATACCAAGTCTCTGACCTTTGACGTCACACATTCAGTCTTG 725  
 QY 241 ThrArgProGluLeuTythrPheValAspValLeuArgSerAlaIleAsnLysThrGlyThr 260  
 Db 726 ACAAGACCAAGACTTACTTGTCTGATGTTCTTCTTCGACATCAACAAACTGGAAGT 785  
 QY 261 AspGluGlyAlaLeuThrArgIleVal 269  
 Db 786 GCCTGAAGAGACATCACTAAGATTGTG 813

RESULT 9  
 CDS0ABV3  
 LOCUS  
 DEFINITION  
 CDS0ABV3 966 bp mRNA linear HTC 06-FEB-2004  
 Arabidopsis thaliana Full-length cDNA Complete sequence from clone  
 GSTRPGH182H06 of Hormone treated Callus of strain col-0 of  
 Arabidopsis thaliana (thale cress).  
 BX815939  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Arabidopsis thaliana (thale cress)  
 Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eustosids II; Brassicales; Brassicaceae; Arabidopsis.  
 1 (bases 1 to 966)  
 Castelli,V., Aury,J.M., Jallion,O., Wincker,P., Clepet,C.,  
 Menard,M., Cruaud,C., Quetier,F., Scarpelli,C., Schachter,V.,  
 Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.  
 Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences:  
 A Combined Approach to Evaluate and Improve Arabidopsis Genome  
 Annotation  
 Unpublished  
 2 (bases 1 to 966)  
 Genoscope.  
 Direct Submission  
 Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :  
 BP 191 91006 Evry cedex - FRANCE (E-mail : seqref@genoscope.cns.fr  
 - Web : www.genoscope.cns.fr)  
 The sequences are based on single pass reads.  
 Life Technologies (a division of Invitrogen) members carried out  
 full-length libraries construction : Temple G.  
 Genoscope members carried out sequencing and annotation : Castelli  
 V., Aury J.M., Jallion O., Wincker P., Menard M., Cruaud C.,  
 Schachter V., Weissenbach J., Salanoubat M.  
 URGV INRA : Clepet C., Caboche M.  
 Annotation is based on the June 2003 version of the Arabidopsis  
 genome released by MIPS (Munich Information center for Protein  
 Sequences). 5 prime and 3 prime are assembled with Phrap.  
 http://www.genoscope.cns.fr/externe/sequences/Banque\_Projet\_EF/Full  
 length  
 http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.  
 1. 966  
 Location/Qualifiers  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /strain="Col-0"  
 /db\_xref="taxon:3702"  
 /clone="GSTRPGH182H06"  
 /tissue\_type="Hormone treated Callus"  
 /plasmid="pCMVSPORT\_6"  
 1. 966  
 /gene="At1g35720"

ORIGIN

Alignment Scores:

Pred. No.: 9,68e-147 Length: 966  
 Score: 1327.00 Matches: 261  
 Percent Similarity: 99.62% Conservative: 1  
 Best Local Similarity: 99.24% Mismatches: 1  
 Query Match: 82.37% Indels: 0

DB: 3 Gaps: 0

US-10-690-564-2 (1-317) x CNS0ABV3 (1-966)

QY 55 GluThrTyrGlyGluAspLeuLeuLysThrLeuAspLysGluLeuSerAsnAspPheGlu 74

Db 2 GAGACCTACGGCGAAGACCTTCTCAAGACTCTTGACAAGAGGCTCTTAAAGATTGGAG 61

QY 75 ArgAlaIleLeuLeuTTPThrLeuGluProGlyGluArgAspAlaLeuLeuAlaAsnGlu 94

Db 62 AGACCTACTTGTGTGGACTCTTGAACCCCGGTGAGCGGTGATGCTTTATTTGGCTTAAGAA 121

QY 95 AlaThrLysArgTTPThrSerSerAsnGlnValLeuMetGluValAlaCysThrArgThr 114

Db 122 GCTTCAAAAAGATGAGCTTCAAGCAACCAAGTCTTTAGAGATTGCTTGACAAGACA 181

QY 115 SerThrglnLeuLeuH1sa1aArgGlnAlaTyrH1sa1aArgTyrLysSerLeuGlu 134

Db 182 TCAACGACAGGTGCTTCAAGCTAAGCAAGCTTACATGCTGCTTCAAGAGTCTTGAA 241

QY 135 GluAspValAlaH1sH1sThrThrGlyAspPheArgLysLeuValSerLeuValThr 154

Db 242 GAGACCTTGGCTCAACCACTACCGGTGACTTGAGAAAGCTTTGGTTCTCTTGTACC 301

QY 155 SerTyrArgTyrGluGluAspGluValAsnMetThrLeuAlaLysGlnGluAlaLysLeu 174

Db 302 TCATACAGGTACGAAGAGATGAAGATGAACATGACATTTGGTGAAGAAAGCTTAAGCTG 361

QY 175 ValH1sGlnLysLysLysAspLysH1sTyrAsnAspGlnAspValIleArgIleLeuSer 194

Db 362 GTCCATGAGAAATCAAGACCAAGCAACACTACATGATGAGATGTTATTAGAAATTTGCC 421

QY 195 ThrArgSerLysAlaGlnIleAsnAlaThrPheAsnArgTyrGlnAspAspH1sGlyGlu 214

Db 422 ACAAGAGCAAAAGCTCAGATCAATGCTTATTTAACTTAACTTAACTTAACTTAACTTAACT 481

QY 215 GluIleLeuLysSerLeuGlnGluGluGluAspAspAspLysPheLeuAlaLeuLeuArg 234

Db 482 GAATTTCCCAAGATCTTGAAGAGAGATGATGATGACAGATCTTGCACTTTAGAG 541

QY 235 SerThrIleGlnCysLeuThrArgProGluLeuTyrPheValAspValIleArgSerAla 254

Db 542 TCAACCATTCAGTCTTGACAAGACCAAGAGCTTACTTGTGATGTTCTTCTGTTACGA 601

QY 255 IleAsnLysThrGlyThrAspGlnGlyAlaLeuThrArgIleValIleThrArgAlaGlu 274

Db 602 ATCAACAAAACTGAGATGATGAAGAGCACTCACTGAATTGGAACCAAGAGCTGAG 661

QY 275 IleAspLeuLysValIleGlyGlnGluTyrGlnArgArgAsnSerIleProLeuGluLys 294

Db 662 ATTGACTTGAAGCTCATTTGAGAGAGATGACCAAGCAAGCAATTTCTTGGAGAAA 721

QY 295 AlaIleThrLysAspThrArgGlyAspTyrGluLysMetLeuValAlaLeuLeuGluGlu 314

Db 722 GCTATTCCAAAGAACTCGTGGAGATTACGAGAAAGTCTCGTGGCACTTCTCGGTGAA 781

QY 315 AspAspAla 317

Db 782 GATGATGCT 790

RESULT 10

LOCUS BU635010 782 bp mRNA linear EST 23-SEP-2002

DEFINITION 003A04 Infected Arabidopsis leaf

ACCESSION sequence.

VERSION BU635010

KEYWORDS BU635010.1 GI:23302265

SOURCE Arabidopsis thaliana (thale cress)

ORGANISM Arabidopsis thaliana

Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE 1 (bases 1 to 782)

AUTHORS Lundsgaard,M., Emmersen,J., Nielsen,K.L., Wilson,I., Somerville,S. and Wellinder,K.G.

TITLE EST sequencing of Erysiphe cichoracearum infected Arabidopsis plants

JOURNAL Unpublished (2002)

COMMENT Contact: Karen G. Wellinder

Institut for Biotechnology

Aalborg Universitet

Søhngårdsøvej 49, 9000 Aalborg, Denmark

Tel: +45 96358467

Fax: +45 98141808

Email: kgyw@bio.auc.dk

FEATURES

source

1. 782

/organism="Arabidopsis thaliana"

/mol\_type="mRNA"

/ecotype="Columbia"

/db\_xref="taxon:3702"

/dev\_stage="Plant 3 weeks old, three days post infection"

/clone\_lib="Infected Arabidopsis Leaf"

/note="Organ: Leaf; Vector: pBluescript; Mixed cDNA library of Arabidopsis and E. cichoracearum infected leaf from three weeks old Arabidopsis plants. Plants were harvested 3 days after infection and mRNA oligo dt selected."

ORIGIN

Alignment Scores:

Pred. No.: 1,31e-138 Length: 782

Score: 1257.50 Matches: 253

Percent Similarity: 99.22% Conservative: 0

Best Local Similarity: 99.22% Mismatches: 2

Query Match: 78.06% Indels: 2

DB: 5 Gaps: 0

US-10-690-564-2 (1-317) x BU635010 (1-782)

QY 1 MetAlaThrLeuLysValSerAspSerValProAlaProSerAspAspAlaGluGlnLeu 20

Db 19 ATGGCAGACCTTTAAAGT-TCGTGATCTGTC-CTGCTCTTCTGATGATGCTGAGCAATTG 76

QY 21 ArgThrAlaPheGlnGlyTTPGlyThrAsnGluAspLeuIleIleSerIleLeuAlaH1s 40

Db 77 AGAACCGCTTTTGAAGATGGGGGTGACGAAGAGCTTATCATCATCTTGGCTCAC 136

QY 41 ArgSerAlaGluGlnArgLysValIleArgGlnAlaTyrH1sGlnThrTyrGlyGluAsp 60

Db 137 AGAAGTCTGACACGAGAAAGTCATCAGGCAAGCATCACGAAACCTACGGCGAAGAC 196

QY 61 LeuLeuLysThrLeuAspLysGluLeuSerAsnAspPheGluArgAlaIleLeuLeuTTP 80

Db 197 CTTCTCAAGACTTGTGACAAGAGCTCTCTAAGATTTCGAGAGACTATCTTGTGTGG 256

QY 81 ThrLeuGluProGlyGluArgAspAlaLeuLeuAlaAsnGluAlaThrLysArgTTPThr 100

Db 257 ACTCTTGAACCCCGGTGAGGTGATGCTTTATTTGGCTTAAGAGCTCAAAAAGATGAGACT 316

QY 101 SerSerAsnGlnValLeuMetGluValAlaCysThrArgThrSerThrglnLeuLeuH1s 120

Db 317 TCAGCAACCAAGTCTTTATGGAAGTTGCTTGACAAGACATCAACGACGCTTTCAC 376

QY 121 AlaArgGlnAlaTyrH1sa1aArgTyrLysSerLeuGlnGluAspValAlaH1sH1s 140

Db 377 GTAGCAAGCTTACCATGCTCGCTACAGAAAGTCTTGAAGAGGACGTTGCTCACAC 436

QY 141 ThrThrglyAspPheArgLysLeuLeuValSerLeuValThrSerTyrArgTyrGlyGlu 160

Db 437 ACTACCGGTGACTCGAAAGCTTTGGTTCTCTTGTTAACCTCATACAGGTACGAAGA 496

QY 161 AspGlnValAsnMetThrLeuAlaLysGlnGluAlaLysLeuValH1sGlnLysLys 180

Db 497 GATGAAGTGAACATGACATTTGGCTAAGCAAGAGCTTAAGCTGTGCTCATGAGAAATCAAG 556

Qy 181 AsplyshistyrAsnaApGluAspValIleArgIleuSerThrArgSerLysAlaGln 200  
 Db 557 GACAAGCACTTACAAATGATGAGATGTTATTAAGATCTTGCCACAAGACCAAGCTCAG 616  
 Qy 201 ILeaenAlaThrPheAsnaArgTyrGlnAspAspHisGlyGlnGluIleuLysSerLeu 220  
 Db 617 ATCAATGCTACTTCTTAAACCGTTACCAAGATGATCATGCGAGGAAATTCACAGAGCTTT 676  
 Qy 221 GluGlnGluAspAspAspAspAspAspAspAspAspAspAspAspAspAspAspAsp 240  
 Db 677 GAGAAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 736  
 Qy 241 ThrArgProGluLeuLysPheValAspValLeuArgSerAlaIle 255  
 Db 737 ACAAGACCGAGCTTACTTGTGAGTGTCTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 781

RESULT 11  
 BU636350 775 bp mRNA linear EST 23-SEP-2002  
 LOCUS 049H05 Infected Arabidopsis leaf Arabidopsis thaliana cDNA, mRNA  
 DEFINITION  
 sequence.  
 ACCESSION BU636350.1 GI:23303605  
 VERSION BU636350.1  
 KEYWORDS EST.  
 SOURCE Arabidopsis thaliana (thale cress)  
 ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
 REFERENCE 1 (bases 1 to 775)  
 AUTHORS Lundgaard,M., Emmersen,D., Nielsen,K.L., Wilson,I., Somerville,S. and Weidner,K.G.  
 TITLE EST sequencing of Erysiphe cichoracearum infected Arabidopsis plants  
 JOURNAL Unpublished (2002)  
 COMMENT Contact: Karen G. Weidner  
 Institut for bioteknologi  
 Aalborg Universitet  
 Sohngaardsholmsvej 49, 9000 Aalborg, Denmark  
 Tel: +45 96358467  
 Fax: +45 98141808  
 Email: kgw@bio.auc.dk.

FEATURES  
 source  
 1..775  
 location/Qualifiers  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /ecotype="Columbia"  
 /db\_xref="taxon:3702"  
 /dev\_stage="Plant 3 weeks old, three days post infection"  
 /clone\_lib="Infected Arabidopsis leaf"  
 /note="Organ: Leaf; Vector: pBluescript; Mixed cDNA library of Arabidopsis and E. cichoracearum infected leaf from three weeks old Arabidopsis plants. Plants were harvested 3 days after infection and mRNA oligo dT selected."

ORIGIN  
 Alignment Score: 3.63e-135 Length: 775  
 Pred. No.: 1228.50 Matches: 249  
 Score: 98.81% Conservative: 1  
 Percent Similarity: 98.42% Mismatches: 2  
 Best Local Similarity: 76.26% Indels: 3  
 Query Match: 5 Gaps: 0  
 DB: US-10-690-564-2 (1-317) x BU636350 (1-775)

Qy 1 MetaLathrLeuLysValSerAspSerValProAlaProSerAspAspAlaGlnLeu 20  
 Db 19 ATGGCGACTTTAAAGG-TCTGATTCGT-CTGCTCTTCTATGATGCTGAGCAATTG 76  
 Qy 21 ArgThrAlaPheGluGlyTTPGlyThrAsnGluAspLeuIleIleSerIleLeuAlaHis 40

Db 77 AGAACCGCTTTGAAGATGGGGTACGAACGAGACTTGATCATATCTTGCTCAC 136  
 Qy 41 ArgSerAlaGluGlnArgValIleArgGlnAlaTyrHisGluThrTyrGluLysP 60  
 Db 137 AGAAGTGTGAACAGAGAAAGTCAATCAGGCAAGATACCAAGAACTTACGGCAAGAC 196  
 Qy 61 LeuLeuLysThrLeuAspLysGluLeuSerAsnaAspPheGluAArgAlaIleLeuLeuTyr 80  
 Db 197 CTTTCTCAAGACTTTCAGACAGAGCTCTTACATTCGATTTCGACAGACTTCTTGTTGCG 256  
 Qy 81 ThrLeuGluProGlyGluArgAspAlaLeuLeuAlaAsnGluAlaThrLysArgTyrThr 100  
 Db 257 ACTCTTGAACCCGGTGAAGCGGTATGCTTATTTGGCTTAATGAAGCTACAAAGATGAGACT 316  
 Qy 101 SerSerAsnGlnValLeuMetGluValAlaCysThrArgThrSerThrGlnLeuLeuHis 120  
 Db 317 TCAAGCAACCAAGTCTTATGGAAGTGTCTTACCAAGGACATCAACGAGCTGCTTCAC 376  
 Qy 121 AlaArgGlnAlaTyrHisAlaArgTyrLysLysSerLeuGluGluAspValAlaHisHis 140  
 Db 377 GCTAGGCAAGCTTACCATGCTGCTTACCAAGAGTCTTTGAAGAGAGACTTGTCTCACAC 436  
 Qy 141 ThrThrGlyAspPheArgLysLeuLeuValSerLeuValThrSerTyrArgTyrGluGly 160  
 Db 437 ACTACCGGTGACTTCAGAAAGCTTTGGTTTCTCTTGTTTACTCATACGATACGAAAGGA 496  
 Qy 161 AspGluValAsnMetThrLeuAlaLysGlnGluAlaLysLeuValHisGluLysLys 180  
 Db 497 GATGAAGTAAACATGACATGTGCTAAGCAAGAGCTAAGCTGTCATGAGAAATTCAG 556  
 Qy 181 AsplyshistyrAsnaApGluAspValIleArgIleLeuSerThrArgSerLysAlaGln 200  
 Db 557 GACAAGCACTTACAAATGATGAGATGTTATTAAGTCTTGCCACAAGACCAAGACTCAG 616  
 Qy 201 ILeaenAlaThrPheAsnaArgTyrGlnAspAspHisGlyGlnGluIleuLysSerLeu 220  
 Db 617 ATCAATGCTACTTCTTAAACCGTTACCAAGATGATCATGCGAGGAAATTCACAGAGCTTT 676  
 Qy 221 GluGlnGluAspAspAspAspAspAspAspAspAspAspAspAspAspAspAspAsp 240  
 Db 677 GAGAAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 736  
 Qy 240 UthraArgProGluLeuLysPheValAspValLeuArg 252  
 Db 737 GACAAGACCGAGCTTACTTGTGAGTGTCTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 773

RESULT 12  
 CB350421 929 bp mRNA linear EST 17-MAY-2003  
 LOCUS P1P02 Cotton fiber subtraced cDNA library Gossypium hirsutum cDNA,  
 DEFINITION  
 mRNA sequence.  
 ACCESSION CB350421  
 VERSION CB350421.1 GI:30841101  
 KEYWORDS EST.  
 SOURCE Gossypium hirsutum (upland cotton)  
 ORGANISM Gossypium hirsutum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.  
 REFERENCE 1 (bases 1 to 929)  
 AUTHORS Li,J., Lu,Y.C., Feng,J.X., Wei,G., Li,J., Shi,Y.H., Fu,Q., Liu,D., Lu,J.C. and Zhu,Y.X.  
 TITLE Isolation and analyses of genes preferentially expressed during early cotton fiber development by subtractive PCR and cDNA array  
 JOURNAL Nucleic Acids Res. 31 (10), 2534-2543 (2003)  
 MEDLINE 22622070  
 PUBMED 12736302  
 COMMENT Contact: Zhu Y.  
 National Laboratory of Protein Engineering and Plant Genetic Engineering of Life Sciences, Peking University  
 Beijing 100871, China

Tel: 86 10 6275 1193  
 Fax: 86 10 6275 4427  
 Email: zhuyxewater@pku.edu.cn.  
 Location/Qualifiers  
 FEATURES  
 source

1. .929  
 /organism="Gossypium hirsutum"  
 /mol\_type="mRNA"  
 /culivar="Xu-143"  
 /db\_xref="taxon:3635"  
 /issue\_type="fiber"  
 /dev\_stage="10 days post anthesis (dpa)"  
 /clone\_lib="Cotton fiber subtracted cDNA library"  
 /note="The library was constructed using PCR-select cDNA subtraction method with 10 dpa cotton fiber as tester and fiberless mutant as driver"

## ORIGIN

## Alignment Scores:

Pred. No.: 2,67e-130 Length: 929  
 Score: 1188.50 Matches: 227  
 Percent Similarity: 85.76% Conservative: 38  
 Best Local Similarity: 73.46% Mismatches: 43  
 Query Match: 73.77% Indels: 1  
 DB: 6 Gaps: 1

US-10-690-564-2 (1-317) x CB350421 (1-929)

QY 1 MetAlaThrLeuLysValSerAspSerValProAlaProSerAspAspAlaGluGlnLeu 20  
 DB 4 ATGGCCACTCTTACAGTGGCCCGACGAGCTTCTTCAAGTCTGAAAGATTGGACCACTA 63  
 QY 21 ArgThrLapheGluGlyTTPGlyThrAsnGlnAspLeuIleIleSerIleLeuAlaHis 40  
 DB 64 AGAAAGCCCTTTTCAGATGGGGAACATGAGGGCTTAATCATATGATATATGGGTGAC 123  
 QY 41 ArgSerAlaGluGlnArgLysValIleArgGlnAlaTyrHisGluThrTyrGlyGluAsp 60  
 DB 124 AGAAATCGGAGCAACGAACTTATTCGAAAACTTACCGTGAACCTTATGGAGAGAT 183  
 QY 61 LeuLeuLysThrLeuAspLysGluLeuSerAsnAspHeGluArgAlaIleLeuLeuTyr 80  
 DB 184 CTCCTCAAGGCACTGGACAAGAGACTCTCGAATGACTTTGAGAGGCGTTCGCTTGG 243  
 QY 81 ThrLeuGluProGlyGluArgAspAlaLeuLeuAlaAsnGluAlaThrLysArgTyrThr 100  
 DB 244 GCTCTTATCTCTGTAACGTGATGCCCTTTGGCTATGATGAGCACAAAGGTGACT 303  
 QY 101 SerSerAsnGlnValLeuMetGluValAlaCysThrArgThrSerThrGlnLeuLeuHis 120  
 DB 304 TCAGACATCATAGTCTTATGAAATAGCTTGCACAAAGTCTGCAACCAACGCTTAC 363  
 QY 121 AlaArgGlnAlaTyrHisAlaArgTyrLysSerLeuGluGluAspValAlaHisHis 140  
 DB 364 GCAAGGAGGCTTATCATGCTCGTTATTAAGAGTCCGTTGAAGAGACGTTGGCTATCAC 423  
 QY 141 ThrThrGlyAspPheArgLysLeuLeuValSerLeuValThrSerTyrArgTyrGlyGly 160  
 DB 424 ACAGCTGGGAGACTTCCATTAAGCTCTCTCTACCTTCTAGTGAATCATCAATATGAGGA 483  
 QY 161 AspGluValAsnMetThrLeuAlaLysGlnGluAlaLysLeuValHisGluLysIleLys 180  
 DB 484 GAGAGAGTGAACATGACTCTGGCAAAACAGAGCGAAGTTCCTCATGAAAAATTTC 543  
 QY 181 AspLysHisTyrAsnAspGluAspValIleArgLysLeuSerThrArgSerLysAlaGln 200  
 DB 544 AACCAAGCTTACAGTATGACGATGATCATTAAGGTTTGGCTCAAGAACCAAGGACAG 603  
 QY 201 IleAsnAlaThrPheAsnArgTyrGlnAspAspHisGlyGluGluIleLeuLysSerLeu 220  
 DB 604 ATCAATGCAACTTAATCACTACAAAAATGAATATGAATGACATTAACCAAGGACTTG 663  
 QY 221 GluGluGlyAspPheAspAspLysPheLeuAlaLeuLeuArgSerThrIleGlnCysLeu 240

DB 664 ---AAGGCTGACCCCTAAGATGAGTTCCTTGCACTACTAAGTCCACAGTGAAGTCTTG 720  
 QY 241 ThrArgProGluLeuTyrPheValAspValLeuArgSerAlaIleAsnLysThrGlyTyr 260  
 DB 721 GTCATCCCGGAAAGATATTTGAAAGGTTCTTCGCTAGCGATCATATACACAGAACG 780  
 QY 261 AspGluGlyValaLeuThrArgAlaIleValThrThrArgAlaGluIleAspLeuValIle 280  
 DB 781 GAATGAAGAGCTCTTTCACAGAGTTGTTGCACTAGGAGGTGAGATTCATTAAGATCAT 840  
 QY 281 GlyGluGlyTyrGlnArgArgAsnSerIleProLeuGluLysAlaIleThrLysAspThr 300  
 DB 841 GCAGATGAGATATGACGAAGAACAGTGTCCCTACTGCTGCTCCATTGTCAAGACACT 900  
 QY 301 ArgGlyAspTyrGlyLysMetLeuVal 309  
 DB 901 CATGAGACTATGAAAAATTGCTGCTG 927

## RESULT 13

COL121757 889 bp mRNA linear EST 16-JUN-2004  
 LOCUS GR\_EB03B10.f GR\_Eb Gossypium raimondii cDNA clone GR\_EB03B10.5',  
 DEFINITION mRNA sequence.  
 ACCESSION COL121757  
 VERSION COL121757.1 GI:48820444  
 KEYWORDS EST.  
 SOURCE Gossypium raimondii  
 ORGANISM Gossypium raimondii  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.

## REFERENCE

## AUTHORS

1 (bases 1 to 889)  
 Kim, H., Yu, Y., Kudrna, D., Hatfield, J., Scum, D., Mueller, C.,  
 Udall, J. A., Rapp, R. A., Wendel, J. F., Rao, K., Soderlund, C. and  
 Wing, R. A.

## TITLE

Global assembly of Cotton ESTs

## JOURNAL

Unpublished (2004)

## COMMENT

Contact: Rod A. Wing  
 Arizona Genomics Institute  
 The University of Arizona  
 Forbes Building Room 303, Tucson, AZ, 85721-0036, USA  
 Tel: 520 626 9595  
 Fax: 520 621 1259  
 Email: http://genome.arizona.edu  
 Plate: 03 row: B column: 10.

## FEATURES

## source

1. .889  
 /organism="Gossypium raimondii"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:29730"  
 /clone="GR\_EB03B10"  
 /issue\_type="floral"  
 /dev\_stage="3 to +3 DPA"  
 /lab\_host="DH10B"  
 /clone\_lib="GR\_Eb"  
 /note="Vector: pCMV.SPORT-6.1; Site 1: NotI; Site 2:  
 EcoRV; library made by invitrogen with RNA supplied by  
 Wendle lab. Directional cloned into NotI-EV. Colonies  
 plated/picked by AGI. More glycerol clones held in -80."

## ORIGIN

## Alignment Scores:

Pred. No.: 2.9e-124 Length: 889  
 Score: 1137.50 Matches: 218  
 Percent Similarity: 86.30% Conservative: 34  
 Best Local Similarity: 74.66% Mismatches: 39  
 Query Match: 70.61% Indels: 1  
 DB: 7 Gaps: 1

US-10-690-564-2 (1-317) x COL121757 (1-889)

QY 1 MetAlaThrLeuLysValSerAspSerValProAlaProSerAspAspAlaGluGlnLeu 20



```

Db      14 ATGGCCACTCTACAGTCCCGCCAGCAGACTTCCTTCAAGTGTCTGAAGATTGTGAACAACATA 73
Qy      21 ArgThrAlaPheGluGlyTTPGlyThrAsnGluAspLeuIleIleSerIleLeuAlaHis 40
Db      74 AGAAAGCCCTTTTCAGAGATGGGAACTAAATGAGGGCTTAATCATAGATATATTGGGTGAC 133
Qy      41 ArgSerAlaGluGlnArgLysValIleArgGlnAlaTyrHisGluThrTyrGlyGluAsp 60
Db      134 AGAAATCCGAGCAGCAACAACTTGATTCCGAAAAACCTACCGCTGAACCTTATGAGAGAGAT 193
Qy      61 LeuLeuLysThrLeuAspLysGluLeuSerAsnAspPheGluArgAlaIleLeuLeuTyr 80
Db      194 CTCCTCAAGGCACATGACAGCAGAGCTCTCCAAATCACTTTGAGGGCTTGCTTCTGCTTTGG 253
Qy      81 ThrLeuGluProGlyGluArgAspAlaLeuLeuAlaAsnGluAlaThrLysArgTTPThr 100
Db      254 GCTCTTGATCTCTGCTGAACGTGATGCCCTTTTGCTTATATAGCCACCAAAAGGTGACT 313
Qy      101 SerSerAsnGlnValLeuMetGluValAlaCysThrArgThrSerThrGlnLeuLeuHis 120
Db      314 TCAAGCAATCAGGTCTTATGAAATAGCCTGCAACAAGTCTGCCAACCAACTGCTTCAC 373
Qy      121 AlaArgGlnAlaTyrHisAlaArgTyrLysLysSerLeuGluAspValAlaHisHis 140
Db      374 GCAAGGAGGCTTATCATGCTCGTTTAAAGAGTCGCTGAAGAGATGTTGCTCATCAC 433
Qy      141 ThrThrGlyAspPheArgLysLeuLeuValSerLeuValThrSerTyrArgTyrGlyGly 160
Db      434 ACGACTGGCCACTTCCTGAAGCTCTCTCTACTTACTGATGATTCATACATATGAGGGA 493
Qy      161 AspGluValAsnMetThrLeuAlaLysGlnGluAlaLysLeuValHisGluLysIleLys 180
Db      494 GAGAGGTGAACATGATCTCGCGCAAAACAGAGCCGAAAGTTCCTCATGAGAAATTTCA 553
Qy      181 AspLysHisTyrAsnAspGluAspValIleArgLysLeuSerThrArgSerLysAlaGln 200
Db      554 GACAAACCTTACATGATGACGATGTCATAGAGGTTTGGCTACAGAAACCAAGGACACAG 613
Qy      201 ILAsnAlaThrPheAsnArgTyrGlnAspAspHisGlyGluGluIleLeuLysSerLeu 220
Db      614 ATCATCAACTGATCTGAATCTACACAGAAATGAATGAAATGACATAAACAGAGACTTG 673
Qy      221 GluGluGlnLysAspAspAspLysPheLeuAlaLeuLeuArgSerThrIleGlnCysLeu 240
Db      674 ---AAGCTGATCTTAAGATGAGATGATTCCTTGCACTAATAAGTCCACAGAGAGTGTG 730
Qy      241 ThrArgProGluLeuTyrPheValAspValLeuArgSerAlaIleAsnLysThrGlyThr 260
Db      731 GTCTATCCGAAAAAGTATTTTGAAGAGTTCCTTCGCTTACCAATCATATAGACGAGAACG 790
Qy      261 AspGluGlnAlaLeuThrArgGluLeuValThrThrArgAlaGluIleAspLeuLysValIle 280
Db      791 GATGAAAGAGCTCTTACTTAAGATGTTGCTTGCACTAGGGCTGAGGTTGATTAAGATCATAT 850
Qy      281 GlyGluGlnTyrGlnArgAsnSerIleProLeu 292
Db      851 GCAGATGAGTACCGCAGAGGAACAGTGTCCCACTG 886

```

```

RESULT 14
LOCUS    COI29429
DEFINITION GR_Eb28E08_r GR_Eb Gossypium raimondii cDNA clone GR_Eb28E08 3',
            mRNA sequence.
ACCESSION COI29429
VERSION    COI29429.1 GI:48878407
KEYWORDS   EST.
SOURCE     Gossypium raimondii
ORGANISM   Gossypium raimondii
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
REFERENCE 1 (bases 1 to 896)
            Kim,H., Yu,Y., Kudrna,D., Hatfield,J., Stum,D., Mueller,C.,

```

```

TITLE
JOURNAL
COMMENT
Udall,J.A., Rapp,R.A., Wendel,J.F., Rao,K., Soderlund,C. and
Wing,R.A.
Global assembly of Cotton ESTs
Unpublished (2004)
Contact: Rod A. Wing
Arizona Genomics Institute
The University of Arizona
Forbes Building Room 303, Tucson, AZ, 85721-0036, USA
Tel: 520 626 9595
Fax: 520 621 1259
Email: http://genome.arizona.edu
Plate: 28 row: E column: 08.
Location/Qualifiers
1..896
source
/organism="Gossypium raimondii"
/mol_type="mRNA"
/db_xref="taxon:29730"
/clone="GR_Eb28E08"
/tissue_type="floral"
/dev_stage="3 to +3 DPA"
/lab_host="DH10B"
/clone_lib="GR_Eb"
/note="vector: pCMV.SPORT-6.1; Site 1: NotI; Site 2:
EcoRV; library made by inviterogen with RNA supplied by
Wendle lab. Directional cloned into NotI-EV. Clones
plated/picked by AGI. More glycerol clones held in -80."
ORIGIN
Alignment Scores:
Pred. No.: 3,61e-121 Length: 896
Score: 1111.50 Matches: 215
Percent Similarity: 83.78% Conservative: 33
Best Local Similarity: 72.64% Mismatches: 47
Query Match: 68.9% Indels: 1
DB: 7 Gaps: 1
US-10-690-564-2 (1-317) x COI29429 (1-896)
Qy      1 MetAlaThrLeuLysValSerAspSerValProAlaProSerAspAspAlaGluGlnLeu 20
Db      9 ATGGCCACTCTTACAGTCCCGCCAGCAGACTTCCTTCAAGTGTCTGAAGATTGTGAACAACATA 68
Qy      21 ArgThrAlaPheGluGlyTTPGlyThrAsnGluAspLeuIleIleSerIleLeuAlaHis 40
Db      69 AGAAACTCTTTTTCAGAGATGGGAACTAAATGAGGGCTTATCATAGATATATTGGGTGAC 128
Qy      41 ArgSerAlaGluGlnArgLysValIleArgGlnAlaTyrHisGluThrTyrGlyGluAsp 60
Db      129 AGAAATCCGAGCAGCAACAACTTGATTCCGAAAAACCTACGTAACCTATGAGAGAGAT 188
Qy      61 LeuLeuLysThrLeuAspLysGluLeuSerAsnAspPheGluArgAlaIleLeuLeuTyr 80
Db      189 CTCCTCAAGGCACATGACAGAGAGCTCTCGAAATGACTTTGAAAGGCTGTGCTTGG 248
Qy      81 ThrLeuGluProGlyGluArgAspAlaLeuLeuAlaAsnGluAlaThrLysArgTTPThr 100
Db      249 GCTCTTGATCTCTGCTGAACGTGATGCCCTTTTGCTTATATAGCCACCAAAAGGTGACT 308
Qy      101 SerSerAsnGlnValLeuMetGluValAlaCysThrArgThrSerThrGlnLeuLeuHis 120
Db      309 TCAAGCAATCAGGTCTTATGAAATAGCCTTACCAAGGTCTGCCAACCAACTGCTTCAC 368
Qy      121 AlaArgGlnAlaTyrHisAlaArgTyrLysLysSerLeuGluAspValAlaHisHis 140
Db      369 GCAAGGAGGCTTATCATGCTCGTTTAAAGATGCTTGAAGAGAGATTTGCTCATCAC 428
Qy      141 ThrThrGlyAspPheArgLysLeuLeuValSerLeuValThrSerTyrArgTyrGlyGly 160
Db      429 ACGACTGGGCACTTCGTAAGTCTCTCTCACTCATAGATGATCATACAGATATGAGGGA 488
Qy      161 AspGluValAsnMetThrLeuAlaLysGlnGluAlaLysLeuValHisGluLysIleLys 180
Db      489 GAGGAGGTGACATGATGCTCGCGCAAAACAGAGGCGAAGTTGCTTCATAGAAAAATTTCA 548

```

QY 181 AspyHstYraenAspGluAspValIleArgIleIeuSerThrArgSerIysAlaGln 200  
 DB 549 GACAAAGCTTACAGTACGAGATGCTTAAGGCTTGGCTTACAAAGAGCAAGGACAG 608  
 QY 201 IleAsnAlaThrPheAsnArgTyrGlnAspAspHisGlyGluGluIleLeuIysSerIeu 220  
 DB 609 ATCAATGCAACTCTGATCTACTACAGATGAATATGGAATGACATTAACAGAGACTTG 668  
 QY 221 GluGluGluAspAspAspAspAspPheLeuAlaLeuLeuArgSerThrIleGlnCysLeu 240  
 DB 669 ---AAGCTGATCTTAAGATGAGATGCTCTTGCCTACTAGGTCACAGTGAAGTCTTG 725  
 QY 241 ThrArgProGluLeuTyrPheValAspValIleuArgSerAlaIleAsnIysThrGlyThr 260  
 DB 726 GTCATATCCGAAAGATATTTGAGAGAGTCTTGCCTAGCAATCAATGACGAGAGACG 785  
 QY 261 AspGluGluAlaLeuThrArgIleValIleThrArgAlaGluIleAspLeuIysValIle 280  
 DB 786 GATTAAGAGAGCTCTTACTAGAGTGTCTTGCACTAGGCTGAGGTTGATCTTAAGATCAT 845  
 QY 281 GlyGluGluTyrGlnArgArgAsnSerIleProLeuGluIysAlaIle 296  
 DB 846 GCAGATGATGACGACGAGAGAAACAGTGTCACTGACTGCTGCTCAT 893  
 RESULT 15  
 COLL6938 881 bp mRNA linear EST 16-JUN-2004  
 LOCUS GR\_EB019H11.r GR\_Eb Gossypium raimondii cDNA clone GR\_EB019H11  
 DEFINITION 3', mRNA sequence.  
 ACCESSION COLL6938  
 VERSION COLL6938.1 GI:48815625  
 KEYWORDS EST.  
 SOURCE Gossypium raimondii  
 ORGANISM Gossypium raimondii  
 Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.  
 1 (bases 1 to 881)  
 Kim,H., Yu,Y., Kudrna,D., Halfield,J., Stum,D., Mueller,C.,  
 Udall,J.A., Rapp,R.A., Wendel,J.F., Rao,K., Soderlund,C. and  
 Wang,R.A.  
 TITLE Global assembly of Cotton ESTs  
 JOURNAL Unpublished (2004)  
 COMMENT Contact: Rod A. Wing  
 Arizona Genomics Institute  
 The University of Arizona  
 Forbes Building Room 303, Tucson, AZ, 85721-0036, USA  
 Tel: 520 626 9595  
 Fax: 520 621 1259  
 Email: http://genome.arizona.edu  
 Plate: 019 row: H column: 11.  
 Location/Qualifiers  
 source  
 1..881  
 /organism="Gossypium raimondii"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:29730"  
 /clone="GR\_EB019H11"  
 /rname="GR\_EB019H11"  
 /rsize\_type="floral"  
 /dev\_stage="3 to +3 DPA"  
 /lab\_host="DH10B"  
 /clone\_lib="GR\_Eb"  
 /note="Vector: pCMV.SPORT-6.1, Site\_1: NotI, Site\_2:  
 EcoRV; Library made by Invitrogen with RNA supplied by  
 Wendle lab. Directional cloned into NotI-EV. Colonies  
 plated/picked by AGI. More glycerol clones held in -80."

Alignment Scores:  
 Pred. No.: 6.1e-121 Length: 881  
 Score: 1109.50 Matches: 212  
 Percent Similarity: 86.32% Conservative: 34  
 Best Local Similarity: 74.39% Mismatches: 38

Query Match: 68.87% Indels: 1  
 DB: 7 Gaps: 1  
 US-10-690-564-2 (1-317) x COLL6938 (1-881)  
 QY 29 ThrAsnGluAspLeuIleIleSerIleLeuAlaHisArgSerAlaGluGlnArgVal 48  
 DB 2 ACTAATGAGGGCTTAATCATATGATATATTCGTCACAGAAATCCAGACAGAACTTG 61  
 QY 49 IleArgGlnAlaTyrHisGluThrTyrGlyGluAspLeuLeuIysThrIleuAspIysGlu 68  
 DB 62 ATTGCAAAAACCTACCGCTGAAACCTATGAGAGAGATCTCTCAAGGACACTAGCAAGAG 121  
 QY 69 LeuSerAsnAspPheGluArgAlaIleLeuLeuTyrThrLeuGluProGluValAsp 88  
 DB 122 CTCTCAATGACCTTGAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 181  
 QY 89 AlaLeuLeuAlaAsnGluAlaThrIysArgTyrThrSerAsnGlnValIleuMetGlu 108  
 DB 182 GCCCTTTTGGCTTAATGAGCCACCAAAAGTGGACTTCAAGCAATGAGTCTTATGAA 241  
 QY 109 ValAlaCysThrArgThrSerThrGlnLeuLeuHisAlaArgGlnAlaTyrHisAlaArg 128  
 DB 242 ATAGCTGACAGAGCTGCGCAACCAAGTCTTCAAGAGGAGGCTTATGATGCTCG 301  
 QY 129 TyrIleuAspSerIeuGluGluAspValAlaHisHisThrGlyAspPheArgIysLeu 148  
 DB 302 TATTAAGAGAGCTGCTTAAGAGAGATGCTGCTATCACACATGCGCACTTCCCTAAGCTTC 361  
 QY 149 LeuValSerIeuValHisSerTyrArgTyrGluGluAspGluValAsnMetThrLeuAla 168  
 DB 362 CTCTACTCTTATGAGTTCATACAGATATGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 421  
 QY 169 TyGluGluAlaIysLeuValHisGluIysIleIysAspIysHisTyrAsnAspGluAsp 188  
 DB 422 AAAACAGAGCGGAAGTGTGCTATGAGAAATTTGACAGAAAGCTTACAGTATGACAT 481  
 QY 189 ValIleArgGluLeuSerThrArgSerIysAlaGlnIleAsnAlaThrPheAsnArgTyr 208  
 DB 482 GTCATATGAGGTTTGGCTCAAGAGCAAGGACAGATCAATGCACTCTGATCACTGAC 541  
 QY 209 GlnAspAspHisGlyGluGluIleLeuIysSerIeuGluGluGluValAspAspAspIys 228  
 DB 542 AAGATGAAATATGAAATGACATTAACAGAGACTTG---AAGCTGATCTTAAGATGAG 598  
 QY 229 PheLeuAlaLeuLeuArgSerThrIleGlnCysLeuThrArgProGluLeuTyrPheVal 248  
 DB 599 TTCCTTGCACTACTAGGTCACAGAGAGTGTGCTATCCGAAAAATTTTGAG 658  
 QY 249 AspValLeuArgSerAlaIleAsnIysThrGlyThrAspGluGluAlaLeuThrArgIle 268  
 DB 659 AAGGTTCTTCCGCTAGCAATCAATAGACGAGAAACGAGTGAAGAGAGCTCTTACTAGAGTT 718  
 QY 269 ValThrThrArgAlaGluIleAspLeuIysValIleGlyGluGluTyrGlnArgAspAsn 288  
 DB 719 GTTTCACATGAGGCTGATGATGATCAATGACATGAGATGACAGAGAGAGAGAG 778  
 QY 289 SerIleProLeuGluIysAlaIleThrIysAspThrArgGlyAspTyrGluIysMetLeu 308  
 DB 779 AGTGTCACACTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 838  
 QY 309 ValAlaLeuLeuGly 313  
 DB 839 CTGCTACTTGACAGA 853

Search completed: August 22, 2005, 12:55:42  
 Job time : 3746 secs

This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 22, 2005, 09:01:41 ; Search time 262 Seconds  
(without alignments)  
1979.770 Million cell updates/sec

Title: US-10-690-564-2

Perfect score: 1611

Sequence: 1 MATLKXSDVPAPSDABOL.....KOTRGDYEKMLVALLGEDDA 317

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO.spool/US10690564/runat.19082005.172308.25317/app.query.fasta\_1.455  
-DB=Issued Patents NA -QFMT=fastap -SUFIX=rfn -MINMATCH=0.1 -LOOFCG=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=humand0.cdi  
-LIST=45 -DOCALLIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFILE=pct -NORM=ext -HEAPSIZE=500 -MAXLEN=200000000  
-USER=US10690564\_@CGN\_1.1.69\_runat.19082005.172308.25317 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -MARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA: \*  
1: /cgn2\_6/prodata/1/ina/5A.COMB.seq: \*  
2: /cgn2\_6/prodata/1/ina/5B.COMB.seq: \*  
3: /cgn2\_6/prodata/1/ina/6A.COMB.seq: \*  
4: /cgn2\_6/prodata/1/ina/6B.COMB.seq: \*  
5: /cgn2\_6/prodata/1/ina/PCBUS.COMB.seq: \*  
6: /cgn2\_6/prodata/1/ina/backfile1.seq: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1171.5	72.7	1293	3	US-09-325-932A-17
2	866.5	53.8	789	3	US-09-325-932A-20
3	573	32.3	704	3	US-09-325-932A-21
4	520.5	32.3	484	3	US-09-325-932A-18
5	472	29.3	1332	4	US-09-949-016-2081
6	472	29.3	1339	4	US-09-023-655-1149
7	472	29.3	1339	4	US-09-949-016-722
8	461.5	28.6	2386	4	US-09-949-016-4604
9	461.5	28.6	2386	4	US-09-949-016-4605
10	451	27.9	1932	4	US-09-949-016-1750
11	450	27.9	2305	3	US-08-526-136-1
12	450	27.9	2311	3	US-08-526-136-3

13	447	27.7	2054	4	US-09-949-016-1199	Sequence 1199, Ap
14	447	27.7	2054	4	US-09-949-016-1200	Sequence 1200, Ap
15	447	27.7	2120	4	US-09-949-016-1187	Sequence 1187, Ap
16	447	27.7	2120	4	US-09-949-016-1198	Sequence 1198, Ap
17	447	27.7	2137	4	US-09-976-594-348	Sequence 348, App
18	425.5	26.4	1301	4	US-09-023-655-1190	Sequence 1190, App
19	425.5	26.4	1301	4	US-09-949-016-4660	Sequence 4660, Ap
20	412.5	25.6	2105	4	US-09-799-451-493	Sequence 463, App
21	398.5	24.7	1362	4	US-09-023-655-1152	Sequence 1152, Ap
22	398.5	24.7	1362	4	US-09-814-915A-101	Sequence 101, App
23	398.5	24.7	1614	4	US-09-949-016-2431	Sequence 1431, Ap
24	393	24.4	1399	4	US-09-023-655-1142	Sequence 1142, Ap
25	393	24.4	1399	4	US-09-949-016-5663	Sequence 5663, Ap
26	390	24.2	957	1	US-08-125-746-4	Sequence 4, App1
27	390	24.2	1460	6	5225537-3	Patent No. 5225537
28	390	24.2	1460	6	5225537-3	Patent No. 5225537
29	390	24.2	1567	1	US-08-125-746-2	Sequence 2, App1
30	390	24.2	1632	4	US-09-949-016-4306	Sequence 4306, Ap
31	389	24.1	981	3	US-09-324-096A-1	Sequence 1, App1
32	389	24.1	981	3	US-09-324-096A-3	Sequence 3, App1
33	389	24.1	981	3	US-09-324-096A-5	Sequence 5, App1
34	389	24.1	981	4	US-09-970-969-1	Sequence 1, App1
35	389	24.1	981	4	US-09-970-969-3	Sequence 3, App1
36	389	24.1	981	4	US-09-970-969-5	Sequence 5, App1
37	389	24.1	1516	4	US-09-919-172-38	Sequence 38, App1
38	384.5	23.9	1660	4	US-09-919-039-393	Sequence 393, App
39	374.5	23.2	1300	4	US-09-949-016-2753	Sequence 2753, Ap
40	374.5	23.2	5300	4	US-09-949-016-14495	Sequence 14495, A
41	370	23.0	2104	4	US-09-270-767-10308	Sequence 10308, A
42	352.5	21.9	1356	4	US-09-010-147B-15	Sequence 15, App1
43	338.5	20.5	1432	4	US-09-799-451-269	Sequence 269, App
44	330	20.5	221	3	US-09-325-932A-19	Sequence 19, App1
45	306	19.0	1793	4	US-09-620-312D-747	Sequence 747, App

#### ALIGNMENTS

RESULT 1  
US-09-325-932A-17  
; Sequence 17, Application US/09325932A  
; Patent No. 6451604  
; GENERAL INFORMATION:  
; APPLICANT: Plim, Barry  
; APPLICANT: Lasham, Annette  
; TITLE OF INVENTION: Compositions affecting programmed cell  
; TITLE OF INVENTION: death and their use in the modification of forestry plant devel.  
; FILE REFERENCE: 1022  
; CURRENT APPLICATION NUMBER: US/09/325,932A  
; NUMBER OF SEQ ID NOS: 206  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 17  
; LENGTH: 1293  
; TYPE: DNA  
; ORGANISM: Pinus radiata  
US-09-325-932A-17

Alignment Scores:  
Pred. No.: 9,44e-139  
Score: 1171.50  
Percent Similarity: 83.608  
Best Local Similarity: 71.298  
Query Match: 72.728  
DB: 3  
Gaps: 1

US-10-690-564-2 (1-317) x US-09-325-932A-17 (1-1293)

QY	1	MetAlaThrLeuLysValSerAspSerValProAlaProSerAspAlaGluGlnLeu	20
DB	95	ANGTCACACTCACCCTCCGACGACGACCTCCCTGTAGCCGATGACGACGACGCTC	154
QY	21	ArgThrAlaPhenGluGlyTrpGlyThrArgGlnLysAspLeuIleIleSerIleLeuAlaHis	40

```
Db      155  CGGACAGCCTTCGAGGATGGGACCAATGAGAGCTGATCATATTCATATTTGGCTCAT 214
Qy      41   ArgSerAlaGluGlnArgValIleArgGlnAlaTyrtHseGlnHrtyrGlyValuAP 60
Db      215  AGGAATGGCGCGGAGAGAACTGATTCGGCAAACTATGCCAGACTTACCGCGAGAGAC 274
Qy      61   LeuLeuYerThrLeuAspLysGluLeuSerAsnAspPheGluuArgAlaIleLeuLeuTrp 80
Db      275  CTCCTCAAGGCATTTGGACAGAACTTACCAATGATTTGGAGAGGCGTGGTCTTTGG 334
Qy      81   ThrLeuGluProGlyGlyValArgAspAlaLeuLeuAlaAsnGluAlaThrIlyAspTrpHr 100
Db      335  TCACCTGATCCGCGTGAACGTGATGCGTACTTGGCGCAATGAAACGAGAAATATGAGACT 394
Qy      101  SerSerAsnGlnValLeuMetGluValAlaCysThrAthrThrSerThrGlnLeuLeuHis 120
Db      395  TCAAGCAACAGGTTTCTCATGAGAAATAGCCTGACGAGGCTCTCCGACGAGTTGCTTAA 454
Qy      121  AlaArgGlnAlaTyrtHseAlaArgTyrtLysSerLeuGluLysValAlaIleHisHis 140
Db      455  GCAAGACAAACATATATATGCGCATTCAGAAAGTCAATGAGAGAGAGCGTCCGTCACAC 514
Qy      141  ThrThrGlyAspPheArgLysLeuLeuValSerLeuValThrSerTyrtArgTyrtGlyGly 160
Db      515  ACAACTGGAGATTTTCTTAAGTTGCTGGTACCTCTTGGGAGCTCTTACCTTAATATGAGA 574
Qy      161  AspGluValAsnMetThrLeuAlaLysGlnGluAlaLysLeuValHisGluLysIleLys 180
Db      575  GATGAGGTGATATGACTTTGGCAAAAGCAGAGCTTAAGTACTCCACGAGAAAGATCTCA 634
Qy      181  AspLysHisTyrtAsnAspGluAspValIleArgIleLeuSerThrArgSerLysAlaGln 200
Db      635  GAGAAAGCTTATGCGCATGAGATCTCATAGAGATTTTGGCTACTGAGAGCAAAAGACAG 694
Qy      201  IleAsnAlaThrPheAsnArgTyrtGlnAspAspHisGlyGluGluIleLeuLysSerLeu 220
Db      695  GTCATATGCTACGCTGATCACTCAAAATGAGTTTGGAAATGATATCAACAAAGGATTTG 754
Qy      221  GluGluGluYAspAspAspLysPheLeuAlaLeuLeuArgSerThrIleGlnCysLeu 240
Db      755  AAAAAT---GATCCAAAAGCGCGTTCTTACTTACTGAGACTACAGTAAAGTGGCTG 811
Qy      241  ThrArgProGluLeuTyrtPheValAspValLeuArgSerAlaIleAsnLysThrGlyThr 260
Db      812  ACTCGCCCTGAGAAAGTATTTGAAAAGTTCTTCGCTTACCATCAATTAAGCAGAGACA 871
Qy      261  AspGluGlyAlaLeuThrArgIleValThrThrArgAlaGluIleAspLeuLysValIle 280
Db      872  GATGAAGGGGCTCTGACCAAGAGTAGTTGCTACAGAGGCCGAGGTTGACATGAAGTTTAA 931
Qy      281  GlyGluGluTyrtGlnArgArgAsnSerIleProLeuGluLysAlaIleThrLysAspThr 300
Db      932  AGTAGAGAGTACCAAGAGAGAAATAGCATCTCTGATGTGCCATTTGCAAGAGAACT 991
Qy      301  ArgGlyAspTyrtGluLysMetLeuValAlaLeuLeuArgSerThrIleGlnCysLeu 317
Db      992  ACTGAGAGCATATGAAAAATGCTTTCGGCATTTGATGGCCAGCTCGAGGCT 1042

RESULT 2
US-09-325-932A-20
; Sequence 20, Application US/09325932A
; Patent No. 6451604
; GENERAL INFORMATION:
; APPLICANT: Flinn, Barry
; TITLE OF INVENTION: Compositions affecting programmed cell
; FILE REFERENCE: 1022
; CURRENT APPLICATION NUMBER: US/09/325,932A
; NUMBER OF SEQ ID NOS: 206
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 20
```

```
; LENGTH: 789
; TYPE: DNA
; ORGANISM: Pinus radiata
US-09-325-932A-20

Alignment Scores:
Pred. No.: 2,58e-100
Score: 866.50
Percent Similarity: 84.17%
Best Local Similarity: 69.58%
Query Match: 53.79%
DB: 3 Gaps: 1

US-10-690-564-2 (1-317) x US-09-325-932A-20 (1-789)

Qy      1  MetAlaThrLeuLysValSerAspSerValProAlaProSerAspAspAlaGluGlnLeu 20
Db      73  ATGGCACTATGCGCGGTGCGACCGTGGTTCGCTCCGCGTGAAGAGATGCCAGACGCTC 132
Qy      21  ArgThrAlaPheGluGlyTyrtPglYthrAsnGluAspLeuIleIleSerIleLeuAlaHis 40
Db      133  CAAAAAGCTTTGGCAGAGATGGGGGACGAATGAATGATCATATCATACTGCTCAC 192
Qy      41  ArgSerAlaGluGlnArgValIleArgGlnAlaTyrtHseGlnHrtyrGlyValuAP 60
Db      193  AGAAACGACGCGCAGCGAAAGTAAATCCGACAAACATATGCCAGACATATGGGAGAGAT 252
Qy      61  LeuLeuYerThrLeuAspLysGluLeuSerAsnAspPheGluuArgAlaIleLeuLeuTrp 80
Db      253  CTTCTCAAGCGCTTGAACAGGAACTCTTATGACTTTGAGAGATCTGTGCTTCTGTGG 312
Qy      81  ThrLeuGluProGlyGlyValArgAspAlaLeuLeuAlaAsnGluAlaThrLysArgTrpHr 100
Db      313  ACCCTGAGATCTCCGAGAGCTGATGATCTTGTCCAAATGAAGCTACCAAGAGATTTGACT 372
Qy      101  SerSerAsnGlnValLeuMetGluValAlaCysThrArgThrSerThrGlnLeuLeuHis 120
Db      373  TCGAGCAATGGGTTCTCATGAGAAATGCTTCAAGAGGCTCTTCAATGAGATTAATTCAG 432
Qy      121  AlaArgGlnAlaTyrtHseAlaArgTyrtLysLysSerLeuGluLysValAlaIleHisHis 140
Db      433  GTGAGCGAGCCTATATCATGCTGTTATGAATCTTTGAAGAAAGACATCCCATATTCAC 492
Qy      141  ThrThrGlyAspPheArgLysLeuLeuValSerLeuValThrSerTyrtArgTyrtGlyGly 160
Db      493  ACTACTGGGGAATTTCCGCAAGCTGCTTGTCTCTGGCAAGTACCTTTCCGATAGAGGG 552
Qy      161  AspGluValAsnMetThrLeuAlaLysGlnGluAlaLysLeuValHisGluLysIleLys 180
Db      553  CCTGAGGTGAACATGACATTTGGCGAGATCAGAGGCTTAAGATCTTCATGAGAAATTCAC 612
Qy      181  AspLysHisTyrtAsnAspGluLysValIleArgIleLeuSerThrArgSerLysAlaGln 200
Db      613  GAGAAAGCTTACCAATCATGATGAGCTCATCAAAATTTGTTATCAAAAGATTAAGCTCAG 672
Qy      201  IleAsnAlaThrPheAsnArgTyrtGlnAspAspHisGlyGluGluIleLeuLysSerLeu 220
Db      673  CTTATGCAACCTCATATCTACCAACATAGAGTTTGGAAATGCCATCAACAAGATCTG 732
Qy      221  GluGluGluYAspAspAspLysPheLeuAlaLeuLeuArgSerThrIleGlnCysLeu 240
Db      733  ---AAGGCTGATCCAAATGATGAATTTCTGAACCTGAGATGAGATTAAGTGGCTTG 789

RESULT 3
US-09-325-932A-21
; Sequence 21, Application US/09325932A
; Patent No. 6451604
; GENERAL INFORMATION:
; APPLICANT: Flinn, Barry
; TITLE OF INVENTION: Compositions affecting programmed cell
; FILE REFERENCE: 1022
```

```
; CURRENT APPLICATION NUMBER: US/09/325,932A
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 206
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 21
; LENGTH: 704
; TYPE: DNA
; ORGANISM: Pinus radiata
US-09-325-932A-21

Alignment Scores:
Pred. No.: 4,43e-63 Length: 704
Score: 573.00 Matches: 109
Percent Similarity: 78.02% Conservative: 33
Best Local Similarity: 59.89% Mismatches: 40
Query Match: 35.57% Indels: 0
DB: 3 Gaps: 0

US-10-690-564-2 (1-317) x US-09-325-932A-21 (1-704)
QY 1 MetAlaThrLeuLySValSerAspSerValProAlaProSerAspAspAlaGluGlnLeu 20
DB 148 ATGTCAACAAATTATAGTCCAGTTCACATACCGACCCCATCTGAAAGCGCTG 207
QY 21 ArgThrAlaPheGluGlyTrpGlyThrAsnGluAspLeuLejIleSerIleLeuAlaHis 40
DB 208 AGAAGAGCTTTGAGGGGTGGGCAAAATAGAAATCAATCAATCAATATTAGACAT 267
QY 41 ArgSerAlaGluGlnArgLySValIleArgGlnAlaTrpHisGluTrpTyGlyGluAsp 60
DB 268 AGAAGTCAACACACACAGCAAAAGTATCCGTCAAGTATTTCACAGTTCGAAAGGAT 327
QY 61 LeuLeuLySThrLeuAspLySgIuLeuSerAsnAspPheGluArgAlaIleLeuLeuTip 80
DB 328 CTCCTTAGAGGATAGATCTGAGCTTTCAGTCACTTTCAGAAAGCTGATTCCTTGG 387
QY 81 ThrLeuGluProGlyIuArgAspAlaLeuLeuAlaAsnGlnAlaThrLySArgTrpThr 100
DB 388 GTACTAGATCCAGTCAAGCTGATGCGGTATCTCATGTGCAATTAAGAAAGTGAAT 447
QY 101 SerSerAsnGlnValIleuMeGluValAlaCysThrArgThrSerThrGlnLeuHis 120
DB 448 GCAGAGATATATCCGCTTTAGAAATTTCCAGTGCCTGATCTTCGCGTGAACATTTGATG 507
QY 121 AlaArgGlnAlaTrpHisAlaArgTyLySLeuSerLeuGluGluAspValAlaHisHis 140
DB 508 GTGAGGCAAGCATATCATATTCCGTACAAAGTCCCTCGAAGAAAGCGTGGTGCACAT 567
QY 141 ThrThrGlyAspPheArgLySLeuLeuValSerLeuValThrSerTyArgTyGlyGly 160
DB 568 ACAAGTGAACCTTCGTAGTGTGCTGAGACCTTGTAAGTTCATATCGTATGAAAGGT 627
QY 161 AspGluValAsnMetThrLeuAlaLySglnGlnAlaLySLeuValHisGluLySilelys 180
DB 628 CCGGAAGTGAATATGATTTGGCAAGTTATGAAAGCAAGAAAGTAAAGTGAATTAATACC 687
QY 181 AspLyS 182
DB 688 GAGCAA 693

RESULT 4
US-09-325-932A-18
; Sequence 18, Application US/09325932A
; Patent No. 6451604
; GENERAL INFORMATION:
; APPLICANT: Plinn, Barry
; APPLICANT: Lasham, Annette
; TITLE OF INVENTION: Compositions affecting programmed cell
; FILE REFERENCE: 1022
; CURRENT APPLICATION NUMBER: US/09/325,932A
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 206
```

```
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 18
; LENGTH: 484
; TYPE: DNA
; ORGANISM: Pinus radiata
US-09-325-932A-18

Alignment Scores:
Pred. No.: 1.16e-56 Length: 484
Score: 520.50 Matches: 106
Percent Similarity: 73.29% Conservative: 12
Best Local Similarity: 65.84% Mismatches: 31
Query Match: 32.31% Indels: 1
DB: 3 Gaps: 1

US-10-690-564-2 (1-317) x US-09-325-932A-18 (1-484)
QY 47 LysValIleArgGlnAlaTrpHisGluTrpTyGlyGluAspLeuLeuLySThrLeuAsp 66
DB 3 AAGCTATTCGGCAAACTATGCGGAGACTTACGGCGAGACCTCTCAAGGCAATTCGAC 62
QY 67 LysGluLeuSerAsnAspPheGluArgAlaIle----- 77
DB 63 AGAAGACTTACCAATGATTTTGAAGT-CTGATCTTCTTAATTATTGTAATTCATCCAT 121
QY 77 ----- 77
DB 122 GGAGACGCGTCCCTCTTCTCAGATTAATCATATTCATTCCTGATGTCAGAGGCTG 181
QY 78 ---LeuLeuTrpThrLeuGluProGlyIuArgAspAlaLeuLeuAlaAsnGlnAlaThr 96
DB 182 GTGGTCTTTGGTGGCTTATCCGCTGAACGATGCGTACTTGGCAATGAAAGGACG 241
QY 97 LysArgTrpThrSerSerAsnGlnValIleuMeGluValAlaCysThrArgThrSerThr 116
DB 242 AAAAGTGAATCTTCAAGCAACAGGTTCTCATGAAATAGCCGTGACAGAGTCCACAG 301
QY 117 GlnLeuLeuHisAlaArgGlnAlaTrpHisAlaArgTyLySLeuSerLeuGluGluAsp 136
DB 302 CAGTCTCATGAGCAAGACATATCATGCTCGATACAAAGATCGCTGAAAGAGGAC 361
QY 137 ValAlaHisHisThrThrGlyAspPheArgLySLeuLeuValSerLeuValThrSerTy 156
DB 362 GTGCTCACACACACACATGAGATTTTGTAAAGTGTGCTGATCTTGTAGCTCTTAC 421
QY 157 ArgTyGlnGlyAspGluValAsnMetThrLeuAlaLySglnGlnAlaLySLeuValHis 176
DB 422 CATTAATGATGAGATGAGTGAATATGACTTTGGCAAAAGCAGAGGCTAAGATCTCAC 481
QY 177 Glu 177
DB 482 GAG 484

RESULT 5
US-09-949-016-2081
; Sequence 2081, Application US/0949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2081
```

```

; LENGTH: 1332
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-2081

Alignment Scores:
Pred. No.:      8.4e-50      Length:      1332
Score:          472.00      Matches:      111
Percent Similarity: 56.86%      Conservative: 63
Best Local Similarity: 36.27%      Mismatches: 130
Query Match:      29.30%      Indels:      2
DB:               4          Gaps:      2

US-10-690-564-2 (1-317) x US-09-949-016-2081 (1-1332)
QY 12 AlaProSerAspAspAlaGluGlnLeuArgThrAlaPheGluGlyTrpGlyThrAsnGlu 31
DB 99 AGCCATCATGAGTGAAGCTATTCAGAAAGCAATCAAGAGAAATTGGAACTGATGAG 158
QY 32 AspLeuIleIleSerIleLeuAlaHisArgSerAlaGluGlnArgValIleArgGln 51
DB 159 AAAATGCTCATGACATCTGACTGAGAGAGGTCAAAATGCACAGCGGCACTGATTGTAAG 218
QY 52 AlaTyHisGluThrTyGlyGluAspLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuSer 71
DB 219 GAATATCAAGCAGCATATGGAAGAGCTGAAAGATGACTTGAAGGGTACTCTCTGCGC 278
QY 72 AspPheGluArgAlaIleLeuLeuTrpThrLeuGluProGlyGluArgAspAlaLeuLeu 91
DB 279 CACTTGGACATCTCATGGTGCGCTTAGTGATCCACCGACAGCTTTGATGCAAAAGCAG 338
QY 92 AlaAsnGluAlaThrIlyArgTrpThrSerSerAsnGlnValLeuMetGluValAlaCys 111
DB 339 CTAAGAGAAATCCATGAAGGCGGGGAGCAACAAGAGATCCCTGATTGAAATCTTAAGT 398
QY 112 ThrArgThrSerThrGlnLeuLeuHisAlaArgGlnAlaTyHisAlaArgTyIlyLeu 131
DB 399 ACCAGAGCAAGCAGCAAAATGAAGAGATATCTCTCAAGCCATATATACAGATACAGAAG 458
QY 132 SerLeuGluGlnArgValAlaHisHisThrThrGlyAspPheArgIlyLeuLeuValSer 151
DB 459 AGCTTGGAGATACATTAAGTCCGAAACATCTGGTGACTTCGGAAGAGCTGTGGTACT 518
QY 152 LeuValThrSerTrpArgTrpGlyGluGlyAspGlnValAsnMetThrLeuAlaIysGln 171
DB 519 TTGCGAGATGCGAGAGAGATGAAGTCTGAAGTGCATGAGATCTGGCCAAACAAGAT 578
QY 172 AlaIlyLeuValHisGluIlyIleIysAspIlyHisIlyTrpAsnArgGluAsp--ValIle 190
DB 579 GCCCAGATTCTCTATAAGCTGGTGGAACAGATGGGGCAGCGATGAAGCAAAATTCAGCT 638
QY 191 ArgIleLeuSerThrArgSerIlyAsnGlnIleAsnAlaThrPheAsnArgTyGlnAsp 210
DB 639 GAGATCCTGTGTTAAGAGCTTCTCTCAATTAAACATAATTTAATGAATACAGAAAT 698
QY 211 AspHisGlyGluGlnIleLeuIlySerLeuGluGlnGlyIlyAspAspAspIlySerPheLeu 230
DB 699 ATGAGCCAAAGAGACATTTGTGCACACATTA--AAAGAGAAATTAATCTGGCAATTTTGA 755
QY 231 AlaLeuLeuArgSerThrIleGlnIlyLeuThrArgProGlnLeuTyTrpPheValAspVal 250
DB 756 GACTTACTGTGGCCATAGTATTATGTGTGAGAAACAGCGCGGCTTTTAAAGCCGAAAGA 815
QY 251 LeuArgSerAlaIleAsnIlyThrGlyThrAspGluGlyAlaLeuThrArgIleValIleTr 270
DB 816 CTGCATCGAGCCCTTGAAGGCTATTTGAAGATGAGATTTACTCTGAACGAAATTAAGGTGG 875
QY 271 ThrArgAlaGluIleAspLeuIlyValIleGlyGluGlyTrpGlnArgArgAsnIle 290
DB 876 TCCAGATCAGAAATTAATCTTTGGACATTCGACACAGAGATTCAGAGAGCATTAAGCTTAT 935
QY 291 ProLeuGluIlyAlaIleThrIlyAspThrArgGlyAspTyGlyIlyIlyMetLeuValAla 310

```

```

DB 936 TCCCTATATTCAGCATTAATCGGATATCTTCTGAGACATTAAGAAATCACTTTAAAA 995
QY 311 LeuLeuGlyGluAspAsp 316
DB 996 ATCTGTGTGAGAGATGAC 1013

RESULT 6
US-09-023-655-1149
; Sequence 1149, Application US/09023655
; Patent No. 6607879
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Susan G. Stuart
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
; TITLE OF INVENTION: EXPRESSION
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/023,655
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0001 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1149:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1339 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: 9186967
; US-09-023-655-1149

Alignment Scores:
Pred. No.:      8.47e-50      Length:      1339
Score:          472.00      Matches:      111
Percent Similarity: 56.86%      Conservative: 63
Best Local Similarity: 36.27%      Mismatches: 130
Query Match:      29.30%      Indels:      2
DB:               4          Gaps:      2

US-10-690-564-2 (1-317) x US-09-023-655-1149 (1-1339)
QY 12 AlaProSerAspAspAlaGluGlnLeuArgThrAlaPheGluGlyTrpGlyThrAsnGlu 31
DB 101 AGCCATCATGAGTGAAGCTATTCAGAAAGCAATCAAGAGAAATTGGAACTGATGAG 160
QY 32 AspLeuIleIleSerIleLeuAlaHisArgSerAlaGluGlnArgValIleArgGln 51
DB 161 AAAATGCTCATGACATCTGACTGAGAGAGGTCAAAATGCACAGCGGCACTGATTGTAAG 220

```

Oy 52 AlaTyrHisGluThrTyrGlyGluAspLeuLeuLysThrLeuAspLysGluLeuSerAsn 71  
Db 221 GAATTCACGACGACATATGAAAGAGAGCTGAAGATGATGAAAGGGATCATCTCTGGC 280			
Oy AspPheGluArgAlaIleLeuLeuThrPheThrLeuGluProGlyGluArgAspAlaLeu 91			
Db 281 CACTTTAGCATCTCATGGTGGCCCTAGTGAATCCACCGACACTCTTTGATGCAAGACAG 340			
Oy 92 AlaAsnGluAlaThrLysArgTyrThrSerSerAsnGlnValLeuMetGluValaLys 111			
Db 341 CTAAGAATAATCATGATGAGGCGCGGGAACAACAGAGATGCCCTTGATTTGAAATCTTAAT 400			
Oy 112 ThrArgThrSerThrGlnLeuLeuHisAlaArgGlnAlaTyrHisAlaArgTyrLys 131			
Db 401 ACCAGACACACAGCGCAATTAAGATATCTCTCAAGCCTAATATACATATACAGAAAG 460			
Oy 132 SerLeuGluGluAspValAlaHisSerThrThrGlyAspPheArgLysLeuLeuValSer 151			
Db 461 AGCTCTTGAGATGACATTGATTCGCCGAACATCTGGTGACTTCGGAAACCTCTGGTGAAT 520			
Oy 152 LeuValThrSerTyrArgTyrGlyAspGlyValAsnMetThrLeuAlaLysGlnGlu 171			
Db 521 TTGGCAGATGCGACGAAGAGATGAAGTCTGAAAGTGAATGAGATCTGGCCCAACAGAT 580			
Oy 172 AlaLysLeuValHisGluLysIleLysAspLysHisTyrAsnAspGluAsp--ValIle 190			
Db 581 GCCCAGATTCCTCTATTAATGCTGAGAACAGATGGGACCGATGATGAACAAATTCACAT 640			
Oy 191 ArgIleLeuSerThrArgSerLysAlaGlnIleAsnAlaThrPheAsnArgTyrGlnAsp 210			
Db 641 GAGATCCCTGTGTTTAAGAGCTTTCCTCAATTAAACCTAATCTATGATGATACAGAAAT 700			
Oy 211 AspHisGlyGluIleLeuLysSerLeuGluGluGlyAspAspAspLysPheLeu 230			
Db 701 ATCAGCCCAAGACATTGTGACACACATA--AAAGGAAATATATCTGGCATTTTGAA 757			
Oy 231 AlaLeuLeuArgSerThrIleGlnCysLeuThrArgProGluLeuTyrPheValAspVal 250			
Db 758 GACTTACTGTGTGACATAGTAAATGTGTGAGAGAACAGCCGGCCTTTTACCGCAAGAA 817			
Oy 251 LeuArgSerAlaIleAsnLysThrGlyThrAspGluGlyAlaLeuThrArgIleValThr 270			
Db 818 CTGCATCAGACCTTGAAGGGTATTTGAACTGATGATTTACTCTGAACCGAATAAAGTGT 877			
Oy 271 ThrArgAlaGluIleAspLeuLysValIleGlyGluGlyThrGlnArgArgAsnSerIle 290			
Db 878 TCCAGATCAGAAATTAATGACCTTTTGGCATTTCGAACAGATCTCAAGAGCATTAATGCTAT 937			
Oy 291 ProLeuGluLysAlaIleThrLysAspThrArgGlyAspTyrGlyLysMetLeuValAla 310			
Db 938 TCCCTATATTCGCAATTAAATCGGATTACTTTCGAGACTATGAATACACACTTTAAAA 997			
Oy 311 LeuLeuGlyGluAspAsp 316			
Db 998 ATCTGTGTGTGAGATGAC 1015			
 RESULT 7  
 ; US-949-016-722  
 ; Sequence 722, Application US/09949016  
 ; Patent No. 6812339  
 ; GENERAL INFORMATION:  
 ; APPLICANT: VENTER, J. Craig et al.  
 ; TITLE OF INVENTION: WITH MORPHISMS IN KNOWN GENES ASSOCIATED  
 ; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
 ; FILE REFERENCE: CLO01307  
 ; CURRENT APPLICATION NUMBER: US/09/949,016  
 ; CURRENT FILING DATE: 2000-04-14  
 ; PRIOR APPLICATION NUMBER: 60/241,755  
 ; PRIOR FILING DATE: 2000-10-20  
 ; PRIOR APPLICATION NUMBER: 60/237,768  
 ; PRIOR FILING DATE: 2000-10-03  
 ; PRIOR APPLICATION NUMBER: 60/231,498  
 ; PRIOR FILING DATE: 2000-09-08  
 ;

```

; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 722
; LENGTH: 1339
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-722

Alignment Scores:
Pred. No.:      8,47e-50      Length:      1339
Score:          472.00      Matches:      111
Percent Similarity: 56.86      Conservative: 130
Best Local Similarity: 36.27      Mismatches:  63
Query Match:      29.30%      Indels:      2
DB:              4          Gaps:      2

US-10-690-564-2 (1-317) x US-09-949-016-722 (1-1339)

OY      12  AAlaProSeRaSpaPaLaGlInLeuArGrThRaLaPheGlInGlyTPRgLYThRaSngLu 31
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB      101 AGCCATCAAGTGTACTGTGAAGCTATTCAGAAAGCAATCAGAGGAATTGCAATGATGAG 160
OY      32  AspLeuLeIleSerLIleLeuLaIhIaRGSerLaGlInaRgLyVaIIleArGIn 51
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB      161 AAATAGTCATCGACATTCGTGACTGAGAGGCTCAATGACACGCGGACTGATGTTAAAG 220
OY      52  ALATyRHiSeGlUnThRyGrLIgLuApeLeuEulYsThRLeuApyLyGluLeuSerAsn 71
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB      221 GAATATCAACGACCATATATGAAAGAGCTGAAAGATGACTGMAAGGGTATCTCTTGGC 280
OY      72  AspPheGlUnARgLaIleLeuLeuTrpThRLeuGInProGlyGIuARgAsPaLaLeuLeu 91
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB      281 CACTTGTGAGCATCTCATGTGTGGCCCTAGTCACTCCACACGACAGTCTTGATGCAAAACAG 340
OY      92  ALAaSnGluLaIThRyLaArGrTrpThRserSerAnGInVaIIleMeCtGluVaIAlaCys 111
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB      341 CTAAAGAAATCCATGAAAGCGCCGGGAAACAAACGAAGTGCCTTGATTTGAAATCTTAAC 400
OY      112 ThARGrThSerThRGrLIeLeuLeuIhIaLaRgInaIATyRHiSaLaRgTyLyLyS 133
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB      401 ACCAGGCAAGCAGGCAATGAAAGATATCTCTCAAGCCATATATACATATACAGAAAG 460
OY      132 SerLeuGluUnARyPaLaIhIaShIaThRGIyaSPheARgLySLeuLeuVaISer 151
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB      461 AGCTTGGAGATGACATTGATTCGGAACATCTGTGTGACTTCGGAAAGCTCTGTACT 520
OY      152 LeuVaIThSerTyARGrTyGrLIgLuYaSPeGluVaIaSnMeCThRLeuAlaLySGLIn 171
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB      521 TTGGCAGATGCGCAAGAGATGAAAGTCTGAAAGATGATGACATCTGGCCCAACAGAT 580
OY      172 ALALyLeuVaIhIeGluLySLeYaSPyShISerYRaNaAPGluApe--ValIle 190
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB      581 GCCACAGATTCTCATTAAGCTGTGTAAGAACAGATGCGGACGAGATGAAGACAAATTCACT 640
OY      191 ArgLIeLeuSerThRGrSerTySaLaGlInIleSnalAThRPhaSnARgTyGrLIaSP 210
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB      641 GAGATCTGTGTTAAGAGACTTCTCTCAATTAACATCAATTAATGATGATACAGAAAT 700
OY      211 AspHiGlyGluGluLIeLeuLySerLeuGluGluGlyaSPaAPaSPaPlyRheLeu 230
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB      701 ATCAGCCAAAGACATGTGTGACAGCATTA--AAAGAGAAATATCTGGGCATTTTGAA 757
OY      231 ALALeuleuARgSerThRIeGInCySeuThRaRGrProGluLeuTyRPhaVaIaSPVaI 256
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB      758 GACTTACTGTGGCCATTAAGTATATGTGTGAGGAACACGCGGCTTTTAAACCGAAAGA 817
OY      251 LeuARgSerLaIleSnulYThRGIyThRaSPGluGlyLaLeuThRaRgIleVaIThr 270
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB      818 CTGCATCGACGCTTGAAGGATTTGGAACATGAAGTTACTCTGAACCGAATATAGTGTG 877
OY      271 ThARgALaGlUnIleSPeLeuVaIIleGlyGluGluTyGrLIaRgARgAenSerIle 290
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB      878 TTCAGATCGAAATTAACCTTTTGACACTTGGAAACAGAGTTCAGAAAGCATTAATGCTAT 937

```





Oy LeuileileSerlleuenaIahisAgsSerlaagluInArblyValIleargInaIa 52  
 Db 1178 AChAATCATGATATATCATCACGACCCGACCAATGTCAGGGGACAGAGATCCGGACGACC 1237  
 Oy 53 TyrhIsGluThrTyrgIyGluAspLeuLeuLySthrLeuAspLySgluSeuSerAAsp 72  
 Db 1238 TTCAAGTCACATTTGGCCGGGACCTTAATGACCTGACCTGAAGCTGTAGATCTCTGGAGAC 1297  
 Oy 73 PhegluArgAlaIleleuLeuTrpThrleuGluProGlyGluArgAspAlaleuLeuAla 92  
 Db 1298 CTGGCAAGGCTGATTCCTGGGGCTCATGATGCCACCGGCCCATTAACGATCCAAAGCAGTTG 1357  
 Oy 93 AsnGluAlaThrLyAspArgTrpThrSerSerAsnGlnValLeuMetGluValAlaCySThr 112  
 Db 1358 AAGAAAGCCATGAGAGGAGCCGACAGATGAAGAAAGCTCTTAATTAATTCCTGGCCACT 1417  
 Oy 113 ArgThrSerThrGlnleuLeuHisAlaIArgGlnAlaTyrrHisAlaArgTyrrLySlySer 132  
 Db 1418 CGACCAATGCTGAAATCCGGGACCATCAATGAAGCCTTAATGAGAGACATTCACAAAGTCC 1477  
 Oy 133 LeuGluGluAspValAlaHisIsthTrhrgIyAspPheArgLyLeuLeuValSerLeu 152  
 Db 1478 CTGGAGAGATGCTCTGAGCTCAGACACATCTGGCCACTTCAGAGAGATCCCATTTCTCTG 1537  
 Oy 153 ValThrSerTyrrArgTyrrGluGluAspGluValAsnMetThrleuAlaIySgInGluAla 172  
 Db 1538 GCCACGGGACATCTGTAGAGAGAGAGAA--AACCTGACCCAGGACAGGGAAAGTCC 1594  
 Oy 173 LySleValAlaHISgluLySlyleuys-----AspLySthTyrrAsn 185  
 Db 1555 CAGGTGGCTGCTGAGATCTTGGAATAGACAGACACCCAGTGGAGACAAAACTTCTCTTG 1654  
 Oy 186 AspGluAspValIleArgIleLeuSerThrArgSerLyAlaGlnIleAsnAlaThrPhe 205  
 Db 1655 GAGACACGTTTCATGACGATCTGTGTACCCGGAGCTATCCGACCTCCGGAGAGCTTC 1714  
 Oy 206 AsnArgTyrr-----GlnAspAspHisgluGluIuleuLySerleuGluGlu 222  
 Db 1715 CAGGAGTTCATCAAGATGACCAACTATGACGTGGAGCACACCACTCAAGAGAGATGTC 1774  
 Oy 223 GlysAspAspAspAspLySthrleuAlaleuLeuArgSerThrIleGlnCySleuThrArg 242  
 Db 1775 GGGAGATGCAGGAGATGCATTTGGGGCATT-----GTTCAAGATGCACAAAGAC 1822  
 Oy 243 ProGluLeuTyrrPheValAspValleuArgSerAlaIleAsnLySthrGlyThrAspGlu 262  
 Db 1823 AAGCTCTCTTCTTGGCCGACAAACTTTCMAATCATGAAGGTCCTGGCACAGATGAG 1882  
 Oy 263 GlysAlaLeuThrArgIleValThrThrArgAlaGlnIleAspLeuLySValIleGlyGlu 282  
 Db 1883 AAGACTCTGACCAAGATCATGTATCCCGCAGAGATGACCTGCTCAACATCCGAGAG 1942  
 Oy 283 GlnTyrrGlnArgArgAsnSerIleProLeuGluLyAlaIleThrLyAspThrArgGly 302  
 Db 1943 GAATTCATTTGAGAAATATGACAAGTCTCTCCACCAAGCATTGAGGGTGACACTCCGGA 2002  
 Oy 303 AspTyrrGluLyMetLeuValAlaLeuLeuGlyGluAspAsp 316  
 Db 2003 GACTTCCTGAAAGCCTGTGCGCTCTCTGTGTGTGTGAGAGAC 2044  
 RESULT 10  
 US-09-949-016-1750  
 ; Sequence 1750, Application US/09949016  
 ; Patent No. 6812339  
 ; GENERAL INFORMATION:  
 ; APPLICANT: VENTER, J. Craig et al.  
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
 ; FILE REFERENCE: CLO01307  
 ; CURRENT APPLICATION NUMBER: US/09/949, 016  
 ; CURRENT FILING DATE: 2000-04-14  
 ; PRIOR APPLICATION NUMBER: 60/241, 755

[illegible]

Db 1470 GAGAGCTCAACAGGCGCATGAGGGGCGAGAAACAAGGACCGGACCTGATTCCATC 1529  
Qy ValThrThrArgAlaGluIleAspLeuValIleGlyGluIleArgGlnArgAsn 288  
Db 1530 ATGTGTCTCGACGACGACGACCTCTCGACATCGATCAAGTAAGCGGATGAC 1589  
Qy 289 SerIleProLeuGluIleValIleThrIleAspThrArgGlyAspTyrGluIleMetLeu 308  
Db 1590 GCGAAGTCGCTGACCAACGACATCTCGGAGATATTCAGGGATTAACCGAAGATTCTG 1649  
Qy 309 ValAlaLeuLeuGlyGluAspAsp 316  
Db 1650 CTGAAGATCTGTGTGGCAATGAC 1673  
RESULT 11  
US-10-526-136-1  
; Sequence 1, Application US/08526136  
; Patent No. 6107089  
; GENERAL INFORMATION:  
; APPLICANT: Towle, Christine A. et al.  
; TITLE OF INVENTION: ANNEXIN XI  
; NUMBER OF SEQUENCES: 36  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: U.S.A.  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
; COMPUTER: IBM PS/2 Model 502 or 55SX  
; OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)  
; SOFTWARE: WordPerfect (Version 5.0)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/526,136  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/214,036  
; FILING DATE:  
; APPLICATION NUMBER: 07/837,775  
; FILING DATE: February 13, 1992  
; APPLICATION NUMBER: 07/764,465  
; FILING DATE: September 23, 1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Clark, Paul T.  
; REGISTRATION NUMBER: 30,162  
; REFERENCE/DOCKET NUMBER: 00786/099001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 542-5070  
; TELEFAX: (617) 542-8906  
; TELEX: 200154  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2305  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; US-08-526-136-1  
Alignment Scores  
Pred. No.: 1,25e-46 Length: 2305  
Score: 450.00 Matches: 111  
Percent Similarity: 54.55% Conservative: 57  
Best Local Similarity: 36.04% Mismatches: 132  
Query Match: 27.93% Indels: 8  
Gaps: 3  
US-10-690-564-2 (1-317) x US-08-526-136-1 (1-2305)  
Qy 13 ProSerAspAspAlaGluIleuArgThrAlaPheGluGlyTyrGlyThrAsnGluAsp 32  
|||||  
:::|||||  
:::|||||

Db 699 CCCCTGCGAGATGCTGAAGTCTCGCGAAGGCCATGAAAGGCTTTGGAGTGAAGACGAG 758  
Qy 33 LeuIleIleSerIleLeuAlaHisArgSerAlaGluGlnArgIleValIleArgGlnAla 52  
Db 759 GCCATATTGACTGCTGCTGGTGTGCTGCTTCAACAAAGAACGACAGAGATCTCTCTGTCG 818  
Qy 53 TyrHisGluThrTyrGlyGluAspLeuLeuIleGlySerIleuAspIleGlyLeuSerAsnAsp 72  
Db 819 TTCAGACACAGCATATGGAAGGATTTGATCAAGAATCTGAATCTGAACTGACAGGAAC 878  
Qy 73 PheGluArgAlaIleLeuLeuThrPheIleGluIleProGlyGluArgAspAlaLeuLeuAla 92  
Db 879 TTTGAGAAACACATCTTGCCCTGATGAAGACCCCTGCTCTTTGACCTTATGAGATA 938  
Qy AenGluAlaThrIleAspArgTrpThrSerAsnGlnValLeuMetGluValAlaCysThr 112  
Db 939 AAGAAAGCTATCAAGGGGGCGGCGACATGATGAAGCTGCTGATGAGATCTCGGCTTCG 998  
Qy 113 ArgThrSerThrGluLeuLeuHisAlaArgGlnAlaTyrHisAlaArgTyrIleIleSer 132  
Db 999 CCGAGCAACGACACATCCGGAGCTGAACAGAGCTTACAGACAGAAATTCAAAAAGACC 1058  
Qy 133 LeuGluGluAspValAlaHisIleThrThrGlyAspPheArgIleLeuLeuValSerLeu 152  
Db 1059 CTGAGAGAGCCATTCGAGCGACACTTCAGGCACTTCACCGGCTCTCATCTCTTC 1118  
Qy 153 ValThrSerTyrArgTyrGluIleAspGluValAsnMetThrLeuAlaIleArgGlnAla 172  
Db 1119 TCTCAGAGGAACCGGATGAAAGCACAAACGTGACATACCTTCTCAGAGAGATGTCG 1178  
Qy 173 LysLeuValHisGluIleLysIleLysAspIleHisIleTyrAsnAspGluAspValIleArg 191  
Db 1179 CAGAGACTGTATGACAGCTGGAGAAACCGCTGGGAACAGATGATGCCAAGTTCAATGCG 1238  
Qy 192 IleLeuSerThrArgSerIleValAlaGlnIleAsnAlaThrPheAsnArgTyrGlnAspAsp 211  
Db 1239 ATTCGTGTCTCCCGAGCCGGCCCACTGTGGCAGTTTAAACGATATCAGAGAGATG 1298  
Qy 212 HisGlyGluGluIleLeuIleSerLeu-----GluGluIleAspAspAspAspLys 228  
Db 1299 ACAGACGTCGACATTGAAAGAGCATCTCGCGGAGATGTCGGGGACCTGAGACAGAGGC 1358  
Qy 229 PheLeuAlaLeuLeuAspSerThrIleGlnCysLeuThrArgProGluLeuTyrPheVal 248  
Db 1359 ATGCTGTGCTGTG-----GTGAATGTCTTAAGAAATACCCAGCCCTTCTTGCT 1406  
Qy 249 AspValLeuArgSerAlaIleAsnIleThrGlyThrAspGluGlyAlaLeuThrArgIle 268  
Db 1407 GAAAGGCTCAACAGGCGCATGAGGGGAGCCGGAACCAAGACCGGACCTGATCCGCATC 1466  
Qy 269 ValThrThrArgAlaGluIleAspLeuValIleGlyGluIleTyrGlnArgArgAsn 288  
Db 1467 ATGTGTCTCGACGACGATTCATCTCTCGACATCAAGCGAGATTAAGCGGCTGTAT 1526  
Qy 289 SerIleProLeuGluIleValIleThrIleAspThrArgGlyAspTyrGlyIleMetLeu 308  
Db 1527 GCGAAGTCGCTGACCAACGACATCAACGGGAGACATTCAGGGATTAACCGAAGATTCTG 1586  
Qy 309 ValAlaLeuLeuGlyGluAspAsp 316  
Db 1587 CTGAAGATCTGTGTGGCAATGAC 1610  
RESULT 12  
US-08-526-136-3  
; Sequence 3, Application US/08526136  
; Patent No. 6107089  
; GENERAL INFORMATION:  
; APPLICANT: Towle, Christine A. et al.  
; TITLE OF INVENTION: ANNEXIN XI  
; NUMBER OF SEQUENCES: 36  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson  
; STREET: 225 Franklin Street

```
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM PS/2 Model 50Z or 55SX
OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
SOFTWARE: Wordperfect (Version 5.0)
CURRENT APPLICATION NUMBER: US/08/526.136
FILING DATE:
APPLICATION NUMBER:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/214.036
FILING DATE:
APPLICATION NUMBER: 07/837,775
FILING DATE: February 13, 1992
APPLICATION NUMBER: 07/764,465
FILING DATE: September 23, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul T.
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 00786/099001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2311
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-526-136-3

Alignment Scores:
Pred. No.: 1.26e-46 Length: 2311
Score: 450.00 Matches: 111
Percent Similarity: 54.55% Conservative: 57
Best Local Similarity: 36.04% Mismatches: 132
Query Match: 27.93% Indels: 8
DB: 3 Gaps: 3

US-10-690-564-2 (1-317) x US-08-526-136-3 (1-2311)
QY 13 ProSeRAspAaPAlaGlulnLeuArgThraAlaPheGlulnGlyTrpGlyThrAsnGluAsp 32
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 705 CCCTGCGAGATGCTGAAGTCCTCCGAGAGCCCATGAAGGCTTTGGGACTGACGACGAG 764
QY 33 LeuIleIleSerIleLeuAlaHsArgSerAlaGlulnArgIysValIleArgGlnAla 52
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 765 GCCATCATTTACTGCTCGGGTATGCTGCCAACACAGACAGACAGCATCTCTGTGCG 824
QY 53 TyrHisGluThrTrpGlyGluAspLeuLeuYThrIleuAspIysGluLeuSerAsnAsp 72
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 825 TTCAAGACACACATATGGGAAGATTTGATCAAGATCTGAATCTGACGAGAAC 884
QY 73 PheGluArgAlaIleLeuLeuTrpThrLeuGluProGlyGluArgAspAlaLeuAla 92
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 885 TTTGAGAGAGACATCTTGGCCCTGATGAAGACCCCTCTCTTGAACGCTTATGAGATA 944
QY 93 AsnGluAlaThrIysValTrpThrSerSerAsnGlnValLeuMetGluValAlaCysThr 112
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 945 AAGGAAGCTATCAAGGGGCGGCGCATGAAAGCCCTGCTGATCGAGATCTGGCCTCC 1004
QY 113 ArgThrSerThrGlnLeuLeuHisAlaArgGlnAlaIleValIleAlaArgTrpIysIysSer 132
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1005 CGGAGCAAGACACATCCGGGAGCTGAACAGACTTACAAAGACAGAAATTCAAAAGACC 1064
QY 133 LeuGluGluAspValAlaIleHisThrTrpGlyAspPheArgIysLeuLeuValSerLeu 152
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1065 CTGGAGGAGGCGCATTCGAGAGACATCTCAAGGGGACTTCAGGGGCTCTCATCTCTCTC 1124
```

```
QY 153 ValThrSerTrpArgTrpGlyGluGluValAsnMetThrIleuAlaIysGlnGluAla 172
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1125 TCTCAGGGAACCGGAGATGAACACAAACGTGGACATGACCTTGTCAGAGAGATGG 1184
QY 173 LysLeuValHisGluIysIleLysAspIysHisIleTyrAsnAspIleAspValIleArg--- 191
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1185 CAGAGACTCTATGACAGTGGGAGAACCCGCTGGGAACAGATAGTCCAAAGTTCATGCG 1244
QY 192 IleuSerThrArgSerIysAlaGlnIleAsnAlaThrPheAsnArgTrpGlnAspAsp 211
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1245 ATTCTGTGCTCCCGAGCCGGGCCACTGTGTGCAGTTTATACAGATTCAGAGATG 1304
QY 212 HisGlyGluGluIleuLysSerLeu-----GluGluGlyAspAspAspIys 228
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1305 ACAGAGCTGACATTAGAAGACATCTCCGAGATGTCCGGGACCTTGAGCAGAGGC 1364
QY 229 PheLeuAlaLeuLeuArgSerThrIleGlnCysLeuThrArgProGluLeuTrpPheVal 248
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1365 ATCTGGCTGTG-----GTGAATGCTTAAAGATACCCAGCCTTCTTGCT 1412
QY 249 AspValLeuArgSerAlaIleAsnIysThrGlyThrAspGluValaLeuThrArgIle 268
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1413 GAAAGGCTCAACAGGCGCATGAGGGAGCCGGAACCAAGACCGGACCTGATCCGCATC 1472
QY 269 ValThrArgAlaGluIleAspLeuIysValIleGlyGluIysTrpGlnArgAsn 288
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1473 ATGTGTCTGCGCAGATTCGACCTCTGGACATGAGCAGATTAAGCGGCTGTAT 1532
QY 289 SerIleProLeuGluIysAlaIleThrIysAspThrArgGlyAspTrpGlyIysMetLeu 308
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1533 GCGAAGTCTGTATCACACACATCACCGGAGACACTTCAGGGGATTAACGGAAATTCG 1592
QY 309 ValAlaLeuLeuGlyGluAspAsp 316
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1593 CTGAGATCTGTGTGCGCATATGC 1616

RESULT 13
US-09-949-016-1199
; Sequence 1199, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTNER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1199
; LENGTH: 2054
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-1199

Alignment Scores:
Pred. No.: 2.51e-46 Length: 2054
Score: 447.00 Matches: 102
Percent Similarity: 54.87% Conservative: 67
Best Local Similarity: 33.12% Mismatches: 121
Query Match: 27.75% Indels: 18
DB: 4 Gaps: 3

US-10-690-564-2 (1-317) x US-09-949-016-1199 (1-2054)
QY 16 AspAlaGluGluIleuArgThrAlaPheGluGlyTrpGlyThrAsnGluAspLeuIleIle 35
```

```

Db      506 GATGAGAAATTCCTGCTAGCAATGCAAGGCTTTTGGACAGATGACGAGCAATGTG 565
Qy      36 Ser1leuLalaHisSerAlaGluGlnArgVal11leArgGlnAlaTyRHisGlu 55
Db      566 GATGTGGTGCCCAACGCTTCATATGATCAGAGCAAAAATTTAAAGCATTTTAAGACC 625
Qy      56 ThrTyG1yGluAspLeuLeuTyThrLeuAspLysGluLeuSerAsnAspPheGluArg 75
Db      626 TCCTATGGCAAGATTTATCAAAAGATCTCAATCAGATTAAGTTGAAATATGGAAGA 665
Qy      76 Ala1leuLeuTyThrLeuGluProGluArgAspAlaLeuLeuAlaGlnGluAla 95
Db      686 CTGATCCTGGCCTTCATGCTCTCACTATTAAGATCCTGAGCTTAACGAAAGCA 745
Qy      96 ThrLysArgTrpHisSerSerAsnGlnValLeuMetGluValAlaCyThrArgTrpSer 115
Db      746 ATCGAGGGAGACAGAACTCAGAAACGTGATTTGATTAAGATTTTGGCAAGAACAAAT 805
Qy      116 ThrGlnLeuLeuHisAlaArgGlnAlaTyRHisAlaArgTyRHisSerLeuGluGlu 135
Db      806 CAGAAATCCGAGAAATTTGCAGATGTTATCAGTCAAGAAATTGGACGAGACCTTGAAAG 865
Qy      136 AspValAlaHisHisThrThrGlyAspPheArgLysLeuLeuValSerLeuVal1ThrSer 155
Db      866 GACATTAGGTCAGATTCATCAGACATTTTGAACGTTTACTGTCTTCATGCTCCAGGGA 925
Qy      156 TyArgTyGluGluArgGluValAsnMetThrLeuAlaLysGlnGluAlaLysLeuVal 175
Db      926 AATCGATAGAAACAGACTATTAACACCAAAATGCTCAGAAATGCTCAGCCTTC 985
Qy      176 HisGluLys1leLysAspLysHisTyRAsnAspGluAspVal1leArg---1leLeuSer 194
Db      986 TATCAAGCTGTGAGGGGAGACCTAGGAGCCGATGAAATCTTGCTTAACATGATCTTGCC 1045
Qy      195 ThrArgSerLysAlaGln1leAsnAlaThrPheAsnArgTyRHisAspHisGluGlu 214
Db      1046 ACAAGAGAGCTTCTCCTAGCTGAGAGCTACCATGAGAGCTTATCTGAGATGCTAATCGA 1105
Qy      215 Glu1leLeuLysSer-----LeuGluGluArgAspAsp 226
Db      1106 GACTGTGATAGCGCTGATGACCCCTGAGTTTCCGATATGTAGAAAGTGT----- 1156
Qy      227 AspLysPheLeuAlaLeuLeuArgSerThr1leGlnCysLeuThrArgProGluLeuTyR 246
Db      1157 -----TTGAGAGACCATCTTGCAAGTGTCCCTGAAACCGCCCTGCTTC 1198
Qy      247 PheValAspValLeuArgSerAla1leAsnLysThrGlyThrAspGluGluAlaLeuThr 266
Db      1199 TTGTCTGAGAGCGCTCTACTATGCTATGAAAGGTGCTGGCAGATGATCCACCCGTGTC 1258
Qy      267 Arg1leVal1ThrThrArgAlaGln1leAspLeuLeuVal1leGluGluGluArgHis 286
Db      1259 CGGATTTGTCTACTGAGAGTGAATGACCTTGTAACAATTAACAGATGTGCTGCTCG 1318
Qy      287 ArgAsnSer1leProLeuGluLysAla1leThrLysAspThrArgLysAspTyRHisGlu 306
Db      1319 ATGTACAGAAAGCTCTGGGCAACAATGATTCAGAGTGACAGAGTGAATTTACCAAGA 1378
Qy      307 MetLeuValAlaLeuLeuGluGlu 314
Db      1379 CTCTCTCTGCTGCTATGTGTGGCCAG 1402

RESULT 14
US-09-949-016-1200
; Sequence 1200, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016

```

```

; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1200
; LENGTH: 2054
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-1200

Alignment Scores:
Pred. No.: 2,51e-46 Length: 2054
Score: 447.00 Matches: 102
Percent Similarity: 54.87% Conservative: 67
Best Local Similarity: 33.12% Mismatches: 121
Query Match: 27,754 Indels: 18
DB: Gaps: 3

US-10-690-564-2 (1-317) x US-09-949-016-1200 (1-2054)
Qy      16 AspAlaGluGlnLeuArgTrpHisAlaPheGluGluTyRArgLysAsnGluAspLeu1le 35
Db      506 GATGAGAAATTCCTGCTAGCAATGCAAGGCTTTTGGACAGATGACGAGCAATGTG 565
Qy      36 Ser1leuLalaHisSerAlaGluGlnArgVal11leArgGlnAlaTyRHisGlu 55
Db      566 GATGTGGTGCCCAACGCTTCATATGATCAGAGCAAAAATTTAAAGCATTTTAAGACC 625
Qy      56 ThrTyG1yGluAspLeuLeuTyThrLeuAspLysGluLeuSerAsnAspPheGluArg 75
Db      626 TCCTATGGCAAGATTTATCAAAAGATCTCAATCAGATTAAGTTGAAATATGGAAGA 685
Qy      76 Ala1leuLeuTyThrLeuGluProGluArgAspAlaLeuLeuAlaGlnGluAla 95
Db      686 CTGATCCTGGCCTTCATGCTCTCACTATTAAGATCCTGAGCTTAACGAAAGCA 745
Qy      96 ThrLysArgTrpHisSerSerAsnGlnValLeuMetGluValAlaCyThrArgTrpSer 115
Db      746 ATCGAGGGAGACAGAACTCAGAAACGTGATTTGATTAAGATTTTGGCAAGAACAAAT 805
Qy      116 ThrGlnLeuLeuHisAlaArgGlnAlaTyRHisAlaArgTyRHisSerLeuGluGlu 135
Db      806 CAGAAATCCGAGAAATTTGCAGATGTTATCAGTCAAGAAATTGGACGAGACCTTGAAAG 865
Qy      136 AspValAlaHisHisThrThrGlyAspPheArgLysLeuLeuValSerLeuVal1ThrSer 155
Db      866 GACATTAGGTCAGATTCATCAGACATTTTGAACGTTTACTGTCTTCATGCTCCAGGGA 925
Qy      156 TyArgTyGluGluArgGluValAsnMetThrLeuAlaLysGlnGluAlaLysLeuVal 175
Db      926 AATCGATAGAAACAGACTATTAACACCAAAATGCTCAGAAATGCTCAGCCTTC 985
Qy      176 HisGluLys1leLysAspLysHisTyRAsnAspGluAspVal1leArg---1leLeuSer 194
Db      986 TATCAAGCTGTGAGGGGAGACCTAGGAGCCGATGAAATCTTGCTTAACATGATCTTGCC 1045
Qy      195 ThrArgSerLysAlaGln1leAsnAlaThrPheAsnArgTyRHisAspHisGluGlu 214
Db      1046 ACAAGAGCTTCTCCTAGCTGAGAGCTACCATGAGAGCTTATCTGAGATGCTAATCGA 1105
Qy      215 Glu1leLeuLysSer-----LeuGluGluArgAspAsp 226
Db      1106 GACTGTGATAGCGCTGATGACCCCTGAGTTTCCGATATGTAGAAAGTGT----- 1156
Qy      227 AspLysPheLeuAlaLeuLeuArgSerThr1leGlnCysLeuThrArgProGluLeuTyR 246
Db      1157 -----TTGAGAGACCATCTTGCAAGTGTCCCTGAAACCGCCCTGCTTC 1198

```



**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 22, 2005, 11:53:35 ; Search time 751 Seconds

(without alignments)  
2742.375 Million cell updates/sec

Title: US-10-690-564-2

Perfect score: 1611  
Sequence: 1 MATLKXSDSVAPASDABDL.....KDRGDEYKMLVALLSDDA 317

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 7316285 seqs, 3248459403 residues

Total number of hits satisfying chosen parameters: 14632570

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=xlh  
-O=/cg2\_1/USPRO.spool/US10690564.rnuc 19082005 172309 25406/app.query.fasta\_1.455  
-DB=Published Applications NA -QFMT=fastap -SUPFIX=rnpb -MINMATCH=0.1  
-LOOPL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR\_SCORE=Pct -THR\_MAX=100  
-THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pico -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-NCPU=6 -ICPU=3 -NO MAP -LARGQUERY -NEG SCORES=0 -MAIT -DSBLCCK=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=10 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-Fgapop=6 -Fgapext=7 -Ygapop=10 -Ygapext=0.5 -Delop=6 -Delext=7

Database : Published Applications NA:

1: /cg2\_6/prodata/2/pubpna/US07\_PUBCOMB.seq:\*  
2: /cg2\_6/prodata/2/pubpna/PCR\_NEW\_PUB.seq:\*  
3: /cg2\_6/prodata/2/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cg2\_6/prodata/2/pubpna/US06\_PUBCOMB.seq:\*  
5: /cg2\_6/prodata/2/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cg2\_6/prodata/2/pubpna/PCRUS\_PUBCOMB.seq:\*  
7: /cg2\_6/prodata/2/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cg2\_6/prodata/2/pubpna/US08\_PUBCOMB.seq:\*  
9: /cg2\_6/prodata/2/pubpna/US09\_PUBCOMB.seq:\*  
10: /cg2\_6/prodata/2/pubpna/US09\_PUBCOMB.seq:\*  
11: /cg2\_6/prodata/2/pubpna/US09C\_PUBCOMB.seq:\*  
12: /cg2\_6/prodata/2/pubpna/US09C\_NEW\_PUB.seq:\*  
13: /cg2\_6/prodata/2/pubpna/US10\_PUBCOMB.seq:\*  
14: /cg2\_6/prodata/2/pubpna/US10\_PUBCOMB.seq:\*  
15: /cg2\_6/prodata/2/pubpna/US10C\_PUBCOMB.seq:\*  
16: /cg2\_6/prodata/2/pubpna/US10C\_PUBCOMB.seq:\*  
17: /cg2\_6/prodata/2/pubpna/US10\_PUBCOMB.seq:\*  
18: /cg2\_6/prodata/2/pubpna/US10C\_PUBCOMB.seq:\*  
19: /cg2\_6/prodata/2/pubpna/US10\_PUBCOMB.seq:\*  
20: /cg2\_6/prodata/2/pubpna/US10\_PUBCOMB.seq:\*  
21: /cg2\_6/prodata/2/pubpna/US10\_PUBCOMB.seq:\*  
22: /cg2\_6/prodata/2/pubpna/US10\_NEW\_PUB.seq:\*  
23: /cg2\_6/prodata/2/pubpna/US11\_PUBCOMB.seq:\*  
24: /cg2\_6/prodata/2/pubpna/US11\_NEW\_PUB.seq:\*  
25: /cg2\_6/prodata/2/pubpna/US60\_NEW\_PUB.seq:\*  
26: /cg2\_6/prodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1611	100.0	1230	21 US-10-690-564-1	Sequence 1, Appl
2	1202.5	74.6	1184	19 US-10-767-795-4242	Sequence 4242, Ap
3	1185.5	73.6	2588	17 US-10-393-840-28	Sequence 28, Appl
4	1177.5	73.1	1145	18 US-10-425-114-7439	Sequence 2938, Ap
5	1177.5	73.1	1145	18 US-10-425-114-29928	Sequence 29928, A
6	1177.5	73.1	1133	18 US-10-424-599-1190	Sequence 1190, Ap
7	1171.5	72.7	1293	14 US-10-219-220-17	Sequence 17, Appl
8	1171.5	72.7	1293	17 US-10-393-840-109	Sequence 109, App
9	1132.5	70.3	1058	14 US-10-219-220-218	Sequence 218, App
10	1112.5	69.1	1145	18 US-10-425-114-10772	Sequence 10772, A
11	1046.5	65.0	1231	18 US-10-424-599-70457	Sequence 70457, A
12	1045	64.9	1321	18 US-10-424-599-70457	Sequence 655, App
13	1011.5	62.8	1385	19 US-10-767-701-13525	Sequence 13525, A
14	1009.5	62.7	1235	18 US-10-425-114-21986	Sequence 21986, A
15	1009.5	62.7	1285	18 US-10-425-114-26803	Sequence 26803, A
16	1009.5	62.7	1972	20 US-10-425-115-2178	Sequence 2178, Ap
17	1008.5	62.6	1235	18 US-10-425-114-35285	Sequence 35285, A
18	1008.5	62.6	1235	18 US-10-425-114-23365	Sequence 23365, A
19	1004.5	62.4	1232	19 US-10-437-963-78307	Sequence 78307, A
20	975.5	60.6	1397	19 US-10-767-701-14545	Sequence 14545, A
21	975.5	60.6	1476	19 US-10-437-963-58743	Sequence 58743, A
22	951.5	58.9	1496	20 US-10-425-115-63693	Sequence 63693, A
23	949.5	58.9	1300	20 US-10-739-930-22726	Sequence 22726, Ap
24	919.5	57.1	908	18 US-10-425-114-10228	Sequence 10228, A
25	866.5	53.8	789	14 US-10-219-220-20	Sequence 20, Appl
26	866.5	53.8	789	17 US-10-393-840-26	Sequence 26, Appl
27	866.5	53.8	789	17 US-10-393-840-112	Sequence 112, Appl
28	832	51.6	1437	18 US-10-424-599-88585	Sequence 88585, A
29	825	51.2	1499	18 US-10-424-599-30062	Sequence 30062, A
30	793	49.2	589	19 US-10-021-323-13465	Sequence 13465, A
31	786.5	48.8	1153	18 US-10-425-114-7980	Sequence 7980, Ap
32	786.5	48.8	1172	18 US-10-425-114-13073	Sequence 13073, A
33	786.5	48.8	2336	18 US-10-424-599-137007	Sequence 137007, A
34	785	48.7	602	19 US-10-021-323-12115	Sequence 12115, A
35	784	48.7	1261	18 US-10-424-599-87476	Sequence 87476, A
36	782	48.5	1194	20 US-10-739-930-3735	Sequence 3735, Ap
37	775	48.1	591	19 US-10-021-323-15046	Sequence 15046, A
38	773.5	47.0	1320	18 US-10-424-599-137005	Sequence 137005, A
39	760	46.2	600	19 US-10-021-323-15621	Sequence 15621, A
40	759	47.1	1113	18 US-10-424-599-130903	Sequence 130903, A
41	757.5	47.0	2132	17 US-10-393-840-27	Sequence 27, Appl
42	742	46.1	618	19 US-10-767-795-4245	Sequence 4245, Ap
43	738	45.8	586	19 US-10-021-323-15526	Sequence 15526, A
44	728	45.2	2612	18 US-10-424-599-68473	Sequence 68473, A
45	691.5	42.9	870	19 US-10-437-963-18738	Sequence 18738, A

#### ALIGNMENTS

RESULT 1  
US-10-690-564-1  
; Sequence 1, Application US/10690564  
; Publication No. US20050089872A1  
; GENERAL INFORMATION:  
; APPLICANT: KOREA KUNHO PETROCHEMICAL CO., LTD  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING ANNEXINS FROM PLANTS  
; FILE REFERENCE:  
; CURRENT APPLICATION NUMBER: US/10/690,564  
; CURRENT FILING DATE: 2003-10-23  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: Kopatentin 1.71  
; SEQ ID NO 1  
; LENGTH: 1230  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana



US-10-690-564-1

## Alignment Scores:

Pred. No.:	1,766-165	Length:	1230
Score:	1611.00	Matches:	317
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	21	Gaps:	0

US-10-690-564-2 (1-317) x US-10-690-564-1 (1-1230)

```
QY      1 MetAlaThrLeuLySerValSerAspSerValProAlaProSerAspAspAlaGluGlnLeu 20
DB      34 ATGGCGACTCTTAAGGTTTCTGATTCTGTTCTCGTCTCTGATGATGCTGAGCAATTG 93
QY      21 ArgThrAlaPheGluGlyTTPGlyThrAsnGluAspLeuLeileSerleleuAlaHis 40
DB      94 AGAACCGCTTTTGAAGATGGGATACGACAGAGACTTGATCATATCAATCTTGCTCAC 153
QY      41 ArgSerAlaGluGlnArgValAlleArgGlnAlaTyHisGluThrTyArgIlyAsp 60
DB      154 AGAAGTCTCTAACAAGAGAAATCATACAGCAAGCATACCAAGAACTTGCGCGAAGAC 213
QY      61 LeuLeuLyThrLeuAspLyGluLeuSerAspPheGluArgAlaIleLeuLeuTTP 80
DB      214 CTTCTCAAGACTCTTGACAGAGAGCTCTTAACGATTTCAGAGAGCTTATCTTGTTGG 273
QY      81 ThrLeuGluProGlyGluArgAspAlaLeuLeuAlaAsnGluAlaThrLyAspTTPThr 100
DB      274 ACTCTTGAAACCGGATGCGTGAATGCTTTATTTGGCTTAATAACTCAAAAAGATGACT 333
QY      101 SerSerAsnGlnValLeuMetGluValAlaCysThrArgThrSerThrGlnLeuLeuHis 120
DB      334 TCAGCACCACCAAGTCTTATGAGAACTTGCTTGCACAGAGCAATCAACGCGCTTCCAC 333
QY      121 AlaArgGlnAlaTyHisAlaArgTyLySerLeuGluGluAspValAlaHisHis 140
DB      394 GTAGGCAAGCTTACCATGCTCGCTACAGAAAGCTCTTGAGAGAGACGTTGCTCACAC 453
QY      141 ThrThrGlyAspPheArgLySerLeuLeuValSerLeuValThrSerTyArgTyArgIly 160
DB      454 ACTACCGGTCATTCAGAAAGCTTTGGTTCTCTGTATCCATACAGAGTCAAGAGGA 513
QY      161 AspGluValAsnMetThrLeuAlaLySerGlnAlaTySerLeuValHisGluLySilelys 180
DB      514 GATGAAGTGAACATGACATTGGCTTAAGCAAGAACTTAAGCTTCCATGAGAAATTCAG 573
QY      181 AspLyHisTyArgAspGluAspValAlleArgIleLeuSerThrArgSerLySalaGln 200
DB      574 GACAAAGCACTACATGATGAGATGTTATTAGATCTTGTCACAAAGAACAAAGCTCAG 633
QY      201 IleAsnAlaThrPheAsnArgTyArgIlyAspAspHisGlyGluGluIleLeuLySerLeu 220
DB      634 ATCAATGCTATCTTTAAACCTTACCAAGATGATCTGCGAGGAAATTCCAAGAGCTTT 693
QY      221 GlnGluGlyAspAspAspAspLyPheLeuAlaLeuLeuArgSerThrIleGlnCysLeu 240
DB      694 GAGGAAGAGATGATGATGACAAAGTCTTGCACTTTTGAGGTCACCAATTCAGTGGCTTG 753
QY      241 ThrArgProGluLeuTyArgPheValAspValLeuArgSerAlaIleAsnLySerGlyThr 260
DB      754 ACAAGACCAAGAGCTTACTTGTGATGTTCTTGCTTGCTGCAATCAACAAAGCTGGAAGT 813
QY      261 AspGluGlyAlaLeuThrArgIleValIleThrThrArgAlaGluIleAspLeuLySalaGln 280
DB      814 GATGAAGAGAGCTCTCAATTAATTTGACCAAGAGCTGAGATTGACTTGAAGGTCATT 873
QY      281 GlnGluGlyThrArgArgAsnSerIleProLeuGlnLySalaIleThrLyAspThr 300
DB      874 GGAAGAGAGATCAAGCGCAAGAACAGCATCTTTGAGAAAGCTTATTACCAAGAGACT 933
QY      301 ArgGlyAspTyArgIlySerMetLeuValAlaLeuLeuGlyGluAspAspAla 317
```

DB 934 CGTGAGATTACAGAAAGATGCTCGTGCACTTCTGATGAGATGATGCT 984

## RESULT 2

US-10-767-795-4242

```

; Sequence 4242, Application US/10767795
; Publication No. US20040181830A1
; GENERAL INFORMATION:
; APPLICANT: Kovalick, David K.
; APPLICANT: Cao, Yongwei
; APPLICANT: Zhou, Yihua
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53534)B
; CURRENT APPLICATION NUMBER: US/10/767,795
; CURRENT FILING DATE: 2004-01-30
; NUMBER OF SEQ ID NOS: 117596
; SEQ ID NO 4242
; LENGTH: 1184
; TYPE: DNA
; ORGANISM: Goosypium hirsutum
; FEATURE:
; OTHER INFORMATION: Clone ID: GOSHI-09MAY01-C101_1
US-10-767-795-4242
```

## Alignment Scores:

Pred. No.:	6,1e-121	Length:	1184
Score:	1202.50	Matches:	230
Percent Similarity:	85.30%	Conservative:	37
Best Local Similarity:	73.48%	Mismatches:	45
Query Match:	74.64%	Indels:	1
DB:	19	Gaps:	1

US-10-690-564-2 (1-317) x US-10-767-795-4242 (1-1184)

```
QY      1 MetAlaThrLeuLySerValSerAspSerValProAlaProSerAspAspAlaGluGlnLeu 20
DB      70 ATGGCCACTTAAACAGTGCACAGCAAGTCTTCCGTCTGAGATGTGAACAGCTTA 129
QY      21 ArgThrAlaPheGluGlyTTPGlyThrAsnGluAspLeuLeileSerleleuAlaHis 40
DB      130 AGAAAGCTTTTTCAGATGGGAACTAATGAGGCTTATATCATATGATATATTTGGCTCAC 189
QY      41 ArgSerAlaGluGlnArgLySerValleArgGlnAlaTyHisGluThrTyArgIlyAsp 60
DB      190 AGAAATGCCAGAGCAAGAACTTGAATGAAAACTTAACGCGAAACCTATGAGAGAGAT 249
QY      61 LeuLeuLyThrLeuAspLyGluLeuSerAspAspPheGluArgAlaIleLeuLeuTTP 80
DB      250 CTCCTCAAGGCACTACAGCAAGAGCTCTCGAAATGACTTTGAGAGGCTGTTTGGCTTGG 309
QY      81 ThrLeuGluProGlyGluArgAspAlaLeuLeuAlaAsnGluAlaThrLyAspTTPThr 100
DB      310 GCTCTTGATTCGCTGAAGCTATGCTTTTGGCTTAATGAAAGCCCAAAAGTGGACT 369
QY      101 SerSerAsnGlnValLeuMetGluValAlaCysThrArgThrSerThrGlnLeuLeuHis 120
DB      370 TCAGCAATCAAGTCTTATGAAATAGCTTGCACAAAGTCTGCCAACCAACTGCTTCCAC 429
QY      121 AlaArgGlnAlaTyHisAlaArgTyLySerLeuGluGluAspValAlaHisHis 140
DB      430 GCAGAGCAAGCTTATCATCTGTTAAAGAGTCCCTGAAGAGATGTTCTCATCAC 489
QY      141 ThrThrGlyAspPheArgLySerLeuValSerLeuValThrSerTyArgTyArgIly 160
DB      490 ACAGCTGGAGACTTCGTAAGCTCTCTTAAGCTTATGATGATGATGATGAGAG 549
QY      161 AspGluValAsnMetThrLeuAlaLySerGlnAlaTySerLeuValHisGluLySilelys 180
DB      550 GAGAGGTGAACATGATGCTGCGAAGAAACAGAGCGAAATTTGCTTCAAGAAAAATTCCA 609
QY      181 AspLyHisTyArgAspGluAspValAlleArgIleLeuSerThrArgSerLySalaGln 200
```

Db 610 GACAAAGCTTACAGTATGACGATGTATTAAGGTTTGGCTACAGAGAACGACAG 669  
Qy 201 ILeasnaIaThrPheAsnaPgyTgIlnAspAphIsgIyGluGluIleuLeuSerleu 220  
Db 670 ATCATATGCACTCTGAAATCACAATAATGAAATATGAAATGAAATGAAATGAAATG 729  
Qy 221 GluIluGluIAspAAspAAspAAspAAspAAspAAspAAspAAspAAspAAspAAsp 240  
Db 730 ---AAGGTGATCTCTAAGATGAGTCTCTTGCACACTAAGCTGACAGAGAGCTGG 786  
Qy 241 ThrArgProGluLeuTyPheValAspValLeuArgSerAlaIleAsnIleThyGlyThr 260  
Db 787 GTCTATCCGAAAAGTATTTTGGAGAGGTTCTTCCGCTGACATCAATAGACGAGAGC 846  
Qy 261 AspGluGluAlaLeuThrArgIleValThrThrArgAlaGluIleAspLeuValIle 280  
Db 847 GATAGAGAGCTCTTACAGATGAGTGTGGACCTAGGCGGTGATCTTAAAGATCAT 906  
Qy 281 GlyGluGluIuTyArgArgArgSerIleProLeuGluValAlaIleThrIleAspThr 300  
Db 907 GCAGATGATACAGACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 966  
Qy 301 ArgGlyAspTyArgIuTyMetLeuValAlaLeuLeuGly 313  
Db 967 CATGAGAGCTATGAAAAATTCCTGCTGCTACTTGCACAGA 1005

RESULT 3  
US-10-393-840-28  
; Sequence 28, Application US/10393840  
; Publication No. US20030229922A1  
; GENERAL INFORMATION:  
; APPLICANT: Bloksberg, Leonard N.  
; TITLE OF INVENTION: Materials and Methods for the  
; FILE REFERENCE: 11000.1012C3  
; CURRENT APPLICATION NUMBER: US/10/393.840  
; PRIORITY FILING DATE: 2003-03-20  
; PRIORITY APPLICATION NUMBER: US 09/636,800  
; PRIORITY FILING DATE: 2000-08-10  
; PRIORITY APPLICATION NUMBER: US 09/170,862  
; PRIORITY FILING DATE: 1998-10-13  
; PRIORITY APPLICATION NUMBER: US 60/148,426  
; PRIORITY FILING DATE: 1999-08-11  
; PRIORITY APPLICATION NUMBER: PCT NZ/99/00169  
; NUMBER OF SEQ ID NOS: 956  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 28  
; LENGTH: 2588  
; TYPE: DNA  
; ORGANISM: Eucalyptus grandis  
US-10-393-840-28

Alignment Scores:  
Pred. No.: 1 3e-118 Length: 2588  
Score: 1185.50 Matches: 228  
Percent Similarity: 83.91% Conservative: 38  
Best Local Similarity: 71.92% Mismatches: 50  
Query Match: 73.59% Indels: 1  
DB: 17 Gaps: 1

US-10-690-564-2 (1-317) x US-10-393-840-28 (1-2588)

Qy 1 MetAlaThrLeuTyValSerAspSerValProAlaProSerAspAAspAlaGluIleu 20  
Db 1390 ATGTCGACTCTCACCGTCCGACACCACTGCCCTGTAGCCGATGATGACGACACTC 1449  
Qy 21 ArgThrAlaPheGluGlyTTPGlyThrAsnGluAspLeuIleIleSerIleLeuAlaHis 40  
Db 1450 CGGACAGCCTTCGACGATGGGAAACAATGAGAGCTGATCAATCATATTCGTCAT 1509  
Qy 41 ArgSerAlaGluIuTyArgValIleArgGluAlaTyThrIleGluIuTyArgIleGluAsp 60

Db 1510 AGGAATGGGGCGGAGAGAGAGTATTCGGCAAACTATGCGGAGACTTACGGCGAGAC 1569  
Qy 61 LeuLeuTyThrLeuAspIuTyLeuSerAsnAspPheGluIuTyAlaIleLeuLeuTyP 80  
Db 1570 CTCCTCAAGCATTGACAGAGAACTTACCAATGATTTCCAGAGGCTGGTGCCTTTGG 1629  
Qy 81 ThrLeuGluProGlyIuTyAspAAspAlaLeuLeuAlaAsnGluIuTyThrIleAspGlyThr 100  
Db 1630 TCCTTGATCGGCTGAACGTGATGCTGACTTGGCCGATATACGAGAGAGAGAGAGAGACT 1689  
Qy 101 SerSerAsnGluIuTyLeuMetGluValAlaCysThrArgThrSerThrGlnLeuLeuHis 120  
Db 1690 TCAAGCAACAGGTTCTTAAGAAATAGCTTGCACAGAGCTCCGCGACAGTTCCTATAG 1749  
Qy 121 AlaArgIuAlaTyThrIleAlaArgTyTyLeuLeuSerLeuGluIuTyValAlaHisHis 140  
Db 1750 GCAAGCAACATATCATGATGCCCATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1809  
Qy 141 ThrThrGlyAspPheArgIuTyLeuLeuValSerLeuValThrSerTyArgTyArgIuTy 160  
Db 1810 ACAACTGAGATTTTGTAAAGTGTGTGATCTTGTGTGAGCTCTTACCGTTATGATGGA 1869  
Qy 161 AspGluValAsnMetThrLeuAlaTyGluGluIuAlaTyLeuValHisGluTyIleTy 180  
Db 1870 GATGAGGTGAATATGACTTTGGCAAAAGCAGAGGCTTAAGTACTCCACGAGAGATCTCA 1929  
Qy 181 AspIleHisTyArgAsnAspGluAspValIleArgIleLeuSerThrArgSerTyAlaGln 200  
Db 1930 GAGAGGCTTATGCGCATGAGATCTCATAGATTTTGGCTCTAGAGCAAGCAAGCAG 1989  
Qy 201 ILeasnaIaThrPheAsnaPgyTgIlnAspAphIsgIyGluGluIleuLeuSerleu 220  
Db 1990 GTCAATGCTTACGGTGAATCATCAAAAAGAGTTGGAATATATATCAACAGAGATTGG 2049  
Qy 221 GluGluGluIAspAAspAAspAAspAAspAAspAAspAAspAAspAAspAAspAAsp 240  
Db 2050 AAAACT---GATCCAAAAGCGCGTCTCTTACTATCTAGAGACTACAGTAAGTGCCTG 2106  
Qy 241 ThrArgProGluLeuTyPheValAspValLeuArgSerAlaIleAsnIleThyGlyThr 260  
Db 2107 ACTGCGCCTGAGAAAGTATTTGAAAAGTTCCTGCTGACCATCAATAGCAGAGAA 2166  
Qy 261 AspGluGluAlaLeuThrArgIleValThrThrArgAlaGluIleAspLeuValIle 280  
Db 2167 GATGAAGGGCTGTGACAGAGATGTTCTTACAGAGGCGGAGGTTGACATGAGATTATA 2226  
Qy 281 GlyGluGluIuTyArgArgArgSerIleProLeuGluValAlaIleThrIleAspThr 300  
Db 2227 AGTGAAGAGTACAGAGAGAGAGATGATCCCTCTGATGCTCCATTGTCAAGAGACT 2286  
Qy 301 ArgGlyAspTyArgIuTyMetLeuValAlaLeuLeuGlyGluAspAAspAla 317  
Db 2287 ACTGAGAGCTATGAAAAATTCCTTGGCATTGATTTGGCAGCTGAGGCT 2337

RESULT 4  
US-10-425-114-7439  
; Sequence 7439, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jindong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(5313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; PRIORITY FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 7439  
; LENGTH: 1145

```

; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700661911_P11
US-10-425-114-7439

```

Alignment Scores:	
Pred. No.:	3, 11e-118
Score:	1177.50
Percent Similarity:	84.54%
Best Local Similarity:	71.29%
Query Match:	73.09%
DB:	18
	Gaps:
	1
	Indels:
	1
	Mismatches:
	48
	Conservative:
	226
	Length:
	1145

US-10-690-564-2 (1-317) x US-10-425-114-7439 (1-1145)

QY	MeAl	Thr	Leu	Val	Ser	Asp	Ser	Val	Pro	Leu	Asp	Asp	Ala	Glu	Gln	Leu	20					
Db	ATGGCGACACTTAAAGGTTCTCTCAGGCTCTTCCCTCCCGGTGCAGATGATTTGGACCACTC	54	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	113					
QY	Arg	Thr	Ala	He	Glu	Glu	Val	Tyr	Phe	Glu	Ala	Leu	Leu	Leu	Leu	Leu	40					
Db	CGAAAAAGCCTTTCAGATGGGGAACTAACGAGAGCTGATGTATGCATCTTGGCTCAT	114	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	173					
QY	Arg	Ser	Ala	Glu	Gln	Val	Leu	Arg	Gln	Ala	Tyr	His	Glu	Val	Thr	Gly	Val	Asp	60			
Db	AGGAATCTGCTCAGAGGAAGGCTTATCAGAGAAAGCTTATGCCAGACCTATGAGGAAGAT	174	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	233				
QY	Leu	Leu	Val	Thr	Leu	Asp	Val	Glu	Ser	Leu	Asp	Asp	Phe	Glu	Ala	Ala	Leu	Leu	Trp	80		
Db	CTCTCTCAAGGCTTTGAGCAAAAGAACTCAGAGATGATTTTGAAGAGGCTGGTTCATCTTTGG	234	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	293			
QY	Thr	Leu	Glu	Val	Pro	Glu	Glu	Val	Ala	Leu	Leu	Ala	Asn	Glu	Ala	Thr	Val	Ser	Thr	100		
Db	ACACTTGAATCTTGCCTTAACGTATGATTTTGGCAAAACGAGGCACATTAATAAATGAGACT	294	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	353			
QY	Ser	Ser	Asn	Gln	Val	Leu	Met	Glu	Val	Ala	Cys	Tyr	Trp	Arg	Thr	Ser	Thr	Gln	Leu	His	120	
Db	TCAGCATATAGGTTCTGGTGAAATATGCTCAGTAAAGTCTCTGAACTATGTTTGGCT	354	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	140		
QY	Ala	Arg	Gln	Ala	Tyr	His	Ala	Arg	Tyr	Val	Ser	Leu	Glu	Val	Ala	His	His	.....	.....	140		
Db	GCGAGGAAGGCTTACCATGATCTTATTAAGAAGTCTCGAGAGAGAGATGTTGCTCATCAC	414	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	473		
QY	Thr	Thr	Glu	Val	Ser	Phe	Arg	Gly	Leu	Leu	Val	Ser	Leu	Val	Thr	Ser	Tyr	Arg	Tyr	Glu	Gly	160
Db	ACAACGAGAGACTTCGTAAAGCTCAATACACCTCTGGTATGTTCTTATCGATTAAGAAAGA	474	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	533	
QY	Asp	Glu	Val	Ala	Asn	Met	Thr	Leu	Ala	Val	Ser	Gln	Ala	Ala	Val	Leu	His	Glu	Val	Leu	His	180
Db	GATGAGGTCACTTGAACCTCGGCAAAAATGAGCAAAATTCGCAAGAAATATTC	534	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	593	
QY	Asp	Val	His	Val	Tyr	Asn	Asp	Glu	Val	Ala	Leu	Arg	Leu	Ser	Thr	Val	Ser	Leu	Val	Gln	200	
Db	AACAAGGCTTATATATATATATACGACCTTCATCAGATTTTGGCCACAGAGACGAGCAAG	594	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	653	
QY	Ala	Asn	Ala	Thr	Thr	He	Asn	Arg	Tyr	Gln	Asn	Asp	Arg	His	Glu	Val	Ser	Leu	220			
Db	ATTAAATCTATCTTTGAATCATCTACAAAGATGATTTGGACAAAGATATACAAAGACCTG	654	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	713	
QY	Glu	Glu	Glu	Glu	Val	Asp	Asp	Asp	Val	Ser	Phe	Leu	Ala	Leu	Leu	Arg	Ser	Thr	Leu	Cys	Leu	240
Db	---AAGGCTATCCAAAGCGAGGTCCTGCTTTATTAAGGCAACTGTGAATGGTTG	714	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	770	
QY	Thr	Arg	Pro	Glu	Leu	Tyr	Phe	Val	Asp	Val	Leu	Arg	Ser	Ala	Leu	Val	Ser	Thr	Gly	Thr	260	
Db	ATCCGTCCCAAGAAAGTACTTCAGAAAGTGTGTTCCCTTGGCGATCAACAAACGAGAAC	771	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	830	
QY	Asp	Glu	Glu	Val	Ala	Leu	Thr	Arg	Glu	Leu	Val	Thr	Trp	Arg	Ala							

Db                   891 GCAGATGAGATCATAAAGAGAGACGTCTCCCTTGAACGGCTATTGTTAAAGACACC 956  
                   :::|||||:::|||||:::|||||:::|||||  
Dy               301 ArgGIAsPDYrGLuYSMeLLeuValAlaleuIeugLyGuASpaPaLa 317  
Db               951 ACCGGACTATGAGAATAATGGCTTGTGGCGCTTTTAGACAATGATGAGCT 1001

RESULT 5  
US-10-42

```

US-10-425-114-29928
Sequence 29928, Application US/10425114
Publication NO. US20040034888A1
GENERAL INFORMATION:
APPLICANT: Liu, Jingdong
APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E
APPLICANT: Tabaska, Jack E
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53313)B
CURRENT APPLICATION NUMBER: US/10/425,114
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ. ID NOS: 73128
SEQ ID NO 29928
LENGTH: 1145
TYPE: DNA
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: UC-GMROPTC033E01_FLI
US-10-425-114-29928

```

Alignment Scores:	
Pred. No.:	3, 1le-118
Score:	1177.50
Percent Similarity:	84.54%
Best Local Similarity:	71.29%
Query Match:	73.09%
DB:	18
Gaps:	1
Length:	1144
Matches:	226
Conservative:	42
Mismatch:	48
Indels:	1
Gaps:	1

US-10-690-564-2 (1-317) x US-10-425-114-29928 (1-1145)

Oy	1	MetAlaThrLeuIleValSerAspSerValProAlaProSerAspAspAlaGluGlnLeu	20
Db	54	ATGGCAACCTTAAGGGTCTTCAGGCTCTCCCCCGTGGAGATGATGGACACTC	113
Oy	21	ArgThrAlaPheGluGlyTyrGlyIleThrAsnGluAspLeuIleIleSerIleLeuAlaHis	40
Db	114	CGAAAGCCTTCAGAGTTGGTGGAACTAAGGAGAGCTGATTTGATGATCTTGGCTCAT	173
Oy	41	ArgSerAlaGluGlnIleArgIleValIleArgGlnAlaTyrHisGluThrTyrGlyValAsp	60
Db	174	AGGAATGCTGCTCAGAGGAAGCTTATCAGAGAACTTAATGCCAGCCTTGGAGAAAT	233
Oy	61	LeuLeuIleThrLeuAspArgGluLeuSerAsnAspPheGluAlaArgAlaIleLeuLeuTyr	80
Db	234	CTCTCAAGGCCCTTGGACAAAGAAATTCACAGATGATTTTGAAGAGCTGGTTCATTTGG	293
Oy	81	ThrLeuGluIleProGluValGluIleArgAspAlaLeuLeuAlaAsnGluAlaThrIleValArgTyrThr	100
Db	294	ACACTTGATTCGTGTAAGCATGATCATTTTGGCAACGAGCGAGCTAAAGAAATGAGACT	353
Oy	101	SerSerAsnGluValLeuMetGluValAlaCysThrArgThrSerThrGlnIleLeuLeuHis	120
Db	354	TCAAGGATCAAGGTTCTGGTGGAAATACCTCGCATAGCTCTTGAAACAAATGTTTGGCT	413
Oy	121	AlaArgGlnAlaTyrHisAlaArgTyrLeuIleYysSerLeuGluGluAspValAlaAlaHisHis	140
Db	414	GCGAGGAAGCTTACATGATCTTTATTAAGAAAGTCTCGAGGAGAGATGGTGGCTCATCAC	473
Oy	141	ThrThrGlyAspPheArgIleGluLeuValSerLeuValThrSerTyrArgTyrGluGly	160
Db	474	ACAACGAGGAGACTTCGTAAAGTCAATGCTACTCTCGGTAACTTTTTCGATTAAGAAAGA	533



```
; SEQ ID NO 17
; LENGTH: 1293
; TYPE: DNA
; ORGANISM: Pinus radiata
US-10-219-220-17

Alignment Scores:
Pred. No.:      1,666-117      Length:      1293
Score:          1171.50      Matches:      226
Percent Similarity: 83.60%      Conservative: 39
Best Local Similarity: 71.29%      Mismatches:  51
Query Match:    72.72%      Indels:      1
DB:              Gaps:      1

US-10-690-564-2 (1-317) x US-10-219-220-17 (1-1293)

QY      1 MetAlaThrLeuLySerAspSerValProAlaProSerAspAlaGluGlnLeu 20
       95 ATGTGCACTCTCACCGCCGCGACCCCTGTCAGTACGATGATCGAGAGCTC 154
       21 ArgThrAlaPheGluGlyTrpGlyThrAsnGluAspLeuLeileSerIleLeuAlaHis 40
       155 CGGACAGCCTTCGACGAGATGGGAAACAATGAGAAGCTGATCATTCATATTGGGTCA 214
       41 ArgSerAlaGluGlnArgLysValIleArgGlnAlaTyrHisGluThrTyrglyGluAsp 60
       215 AGGAATGGCGCGCAGAGGAAGCTGATTCCGCAAACTATGCCAGACTTACGGCGAGAC 274
       61 LeuLeuLyThrLeuAspLysGluLeuSerAsnAspPheGluArgAlaIleLeuLeuTrp 80
       275 CTCCTCAAGGCATTTGACAGAGAATTACCAATGATTTCCAGAGGCTGGTGTCTTTGG 334
       81 ThrLeuGluProGlyGluArgAspAlaLeuLeuAlaAsnGluAlaThrLysArgTrpThr 100
       335 TCACCTTGATCCGCGTGAACGTGATGCTTACTTGCGAATACGACGAAAGAATGGACT 394
       101 SerSerAsnGlnValLeuMetGluValAlaCysThrArgThrSerThrGlnLeuLeuHis 120
       395 TCAGAGCAACCAAGGTTCATGAGAAATAGCTTGACGAGGCTCCGCGACGAGTTGCTTAA 454
       121 AlaArgGlnAlaTyrHisAlaArgTyrLysLysSerLeuGluGluAspValAlaHisHis 140
       455 GCAAGACCAACATATCATGCTCCGATACAGAGAATGATGAGAAGGACGCTGCACAC 514
       141 ThrThrGlyAspPheArgLysLeuLeuValSerLeuValThrSerTyrArgTyrGluGly 160
       515 ACAACTGGAGATTTTGTGAATGTGCTGGTCTCTTGCGAGCTCTACCGTAATAGTGA 574
       161 AspGluValAsnMetThrLeuAlaLysGlnGluAlaLysLeuValHisGluLysIleLys 180
       575 GATGAGGTGAATATGACTTTGGCAAAACGAGGCTTAAGTACTCCACGGAAGATCTCA 634
       181 AspLysHisTyrAsnAspGluAspValIleArgIleLeuSerThrArgSerLysAlaGln 200
       635 GAGAAGGCTTATGCGCCATGAGATCTCATAGGATTTTGGCTACGAGCAAAAGACAG 694
       201 ILAsnAlaThrPheAsnArgTyrGlnAspAspHisGlyGluGluLeuLysSerLeu 220
       695 GTCAATGCTACGCGTAATCTACATAAAATGAGTTTGGAAATAGATATCAACAAAGATT 754
       221 GluGluGlyAspAspAspAspLysPheLeuAlaLeuLeuArgSerThrIleGlnCysLeu 240
       755 AAAAAT---GATCAAAAGAAGCGGTTCTTACTATATCTAGAGCTACAGTAAGTGCGTG 811
       241 ThrArgProGluLeuTyrPheValAspValLeuArgSerAlaIleAsnLysThrGlyThr 260
       812 ACTCGCCCTGAGAGATTTTGAAGAAGTCTTGTGTACCCATCAATAAGCAAGAGAAC 871
       261 AspGluGlyAlaLeuThrArgIleValThrThrArgAlaGluIleAspLeuLysValIle 280
       872 GATGAAGGGCTCTGACCAAGATAGTGTCTACCAAGGCGCAGAGGTGACATGAAGTTTAA 931
       281 GlyGluGlyTyrGlnArgArgAsnSerIleProLeuGluLysAlaIleThrLysAspThr 300
```

```
DB      932 AGTGAAGATACACAGAGGAGATAGCATCCCTCTGCATCGGCACTTGTCAAGACACT 991
QY      301 ArgGlyAspTyrGluLysMetLeuValAlaLeuLeuGlyGluAspAla 317
       992 ACTGAGACTATGAAAAAATGCTTCTGCGATGATGGCAGCTGAGGCT 1042

RESULT 8
US-10-393-840-109
; Sequence 109, Application US/10393840
; Publication No. US20030229922A1
; GENERAL INFORMATION:
; APPLICANT: Blockberg, Leonard N.
; TITLE OF INVENTION: Materials and Methods for the
; FILE REFERENCE: 11000.1012c3
; CURRENT APPLICATION NUMBER: US/10/393,840
; PRIOR FILING DATE: 2003-03-20
; PRIOR APPLICATION NUMBER: US 09/636,800
; PRIOR FILING DATE: 2000-08-10
; PRIOR APPLICATION NUMBER: US 09/170,862
; PRIOR FILING DATE: 1998-10-13
; PRIOR APPLICATION NUMBER: US 60/148,426
; PRIOR FILING DATE: 1999-08-11
; PRIOR APPLICATION NUMBER: PCT NZ/99/00169
; PRIOR FILING DATE: 1999-10-08
; NUMBER OF SEQ ID NOS: 956
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 109
; LENGTH: 1293
; TYPE: DNA
; ORGANISM: Pinus radiata
US-10-393-840-109

Alignment Scores:
Pred. No.:      1,666-117      Length:      1293
Score:          1171.50      Matches:      226
Percent Similarity: 83.60%      Conservative: 39
Best Local Similarity: 71.29%      Mismatches:  51
Query Match:    72.72%      Indels:      1
DB:              Gaps:      1

US-10-690-564-2 (1-317) x US-10-393-840-109 (1-1293)

QY      1 MetAlaThrLeuLySerAspSerValProAlaProSerAspAlaGluGlnLeu 20
       95 ATGTGCACTCTCACCGCCGCGACCCCTGTCAGTACGATGATCGAGAGCTC 154
       21 ArgThrAlaPheGluGlyTrpGlyThrAsnGluAspLeuLeileSerIleLeuAlaHis 40
       155 CGGACAGCCTTCGACGAGATGGGAAACAATGAGAAGCTGATCATTCATATTGGGTCA 214
       41 ArgSerAlaGluGlnArgLysValIleArgGlnAlaTyrHisGluThrTyrglyGluAsp 60
       215 AGGAATGGCGCGCAGAGGAAGCTGATTCCGCAAACTATGCCAGACTTACGGCGAGAC 274
       61 LeuLeuLyThrLeuAspLysGluLeuSerAsnAspPheGluArgAlaIleLeuLeuTrp 80
       275 CTCCTCAAGGCATTTGACAGAGAATTACCAATGATTTCCAGAGGCTGGTGTCTTTGG 334
       81 ThrLeuGluProGlyGluArgAspAlaLeuLeuAlaAsnGluAlaThrLysArgTrpThr 100
       335 TCACCTTGATCCGCGTGAACGTGATGCTTACTTGCGAATACGACGAAAGAATGGACT 394
       101 SerSerAsnGlnValLeuMetGluValAlaCysThrArgThrSerThrGlnLeuLeuHis 120
       395 TCAGAGCAACCAAGTTCATGAGAAATAGCTTGACGAGAGCTTCCGCGACGAGTTGCTTAA 454
       121 AlaArgGlnAlaTyrHisAlaArgTyrLysLysSerLeuGluGluAspValAlaHisHis 140
       455 GCAAGACCAACATATCATGCTCCGATACAGAGAATGATGAGAAGGACGCTGCACAC 514
       141 ThrThrGlyAspPheArgLysLeuLeuValSerLeuValThrSerTyrArgTyrGluGly 160
```

```
|||||
Db 515 ACAACGTGAGATTTTCGTAAGTTCGTGATCTTCTGGAGACTCTTACCGTAAGATGGA 574
Qy 161 AepGlValaIaenMetThrLeuAlaIySgInuAlaIySLeuValIhIsgIuIySIlElys 180
Db 575 GATAGAGGTGAATATGACTTTGGCAAAAGCAGAGCTTAAGATATCTCCAGAGAGATCTCA 634
Qy 181 AspLyshIstYrAenAepGlValaIaIeArGIleuSerThrAgsSerIySAlaIeIn 200
Db 635 GAGAAAGCTTATGGCCATGAGATCTCATAGAGATTTTGGCTACTAGAGCAAGACAG 694
Qy 201 ILeaSnAlaThrPheaSnArGIyRGIhAspAspHIsGIyGluGluIleuIySerIeu 220
Db 695 GTCAATGCTACGCTGAATCACTACAAAATAGATTTGGAATATATCAACAGATTTG 754
Qy 221 GluGIuGIuIyAspAspAspIySPhelEuAlaIeuleuArGIserThriIeGIuCySleu 240
Db 755 AAAACT---GATCCAAAGAGCGCTTCTTACTATACGTAGAGCTACAGTAAAGTGCCTG 811
Qy 241 ThrArpProGIuIeUyRrPheValaIaSpValIeUySeraIaIleAnIySThrGIyThr 260
Db 812 ACTGCGCTGAGAGATTTTGGAAAGGTTCTTGTGTAGCCATCAATTAAGCCAGAGACA 871
Qy 261 AspGIuGIuIyAlaIeUyThrArGIleValIThrThArGIaGIuIleAepIySValIle 280
Db 872 GATGAAGGGGCTCTGACCAAGATGTTGCTTACAGGCGGAGTTGACATGAAGTTTATA 931
Qy 281 GIyGIuGIuIyRGIuArGIaRGIaSenSerIeProIeUySAlaIleThrIySAspThr 300
Db 932 AGTGAAGAGTACCAAGAGAGATAGCATCCCTCTCGATCGTGCATTTGTCAAAGACACT 991
Qy 301 ArgGIyAspTYrGIuIySmetIeUyAlaIaIeUleuGIyGluAspAspAla 317
Db 992 ACTGAGACATATGAAAAAATGCTTGTGCATATGATGGCCACGTGAGGCT 1042
```

## RESULT 9

```
US-10-219-220-218
; Sequence 218, Application US/10219220
; Publication No. US20030082724A1
; GENERAL INFORMATION:
; APPLICANT: Flinn, Barry
; APPLICANT: Lafram, Annette
; TITLE OF INVENTION: Compositions affecting programmed cell
; FILE REFERENCE: death and their use in the modification of plant development
; CURRENT FILING DATE: 11000.1022ct
; PRIOR APPLICATION NUMBER: US/10/219, 220
; PRIOR FILING DATE: 2002-08-14
; PRIOR APPLICATION NUMBER: U.S. No. US20030082724A1 09/325, 932
; NUMBER OF SEQ ID NOS: 290
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 218
; LENGTH: 1058
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
US-10-219-220-218
```

## Alignment Scores:

```
Pred. No.: 2,26e-113 Length: 1058
Score: 1132.50 Matches: 218
Percent Similarity: 83.91% Conservative: 48
Best Local Similarity: 68.77% Mismatches: 50
Query Match: 70.30% Indels: 1
Gaps: 1
```

US-10-690-564-2 (1-317) x US-10-219-220-218 (1-1058)

```
Qy 1 MetAlaThrLeuIySValIaSerSeraIaProAlaProSerAepAlaGIuGIuIleu 20
Db 72 ATGGCGCATATCGGGTGCACCGCTCGTTCGGCTCGGCTGAGAGTGGCGGACACTC 131
Qy 21 ArgThralaPheGIuGIyTPGIyThrAsnGIuAspLeuIleIeSerIleuAlaIhIs 40
```

```
Db 132 CAAAAAGCTTCCGAGATGGGAGACGAATGAAGATCTGATTCATTCACACTGCGCTAC 191
Qy 41 ArgSerAlaGIuGIuArGIySValIleArGIaIaIyRHIsgIuThrYrGIyGIuAsp 60
Db 192 AGAAACGACAGCGAGCGAAAGTAATCCGAAACATATGCTCCGACATATGAGGAAAGAT 251
Qy 61 LeuIleuIyThrIeUyAepIySGIuIeUySerAepAspPheGIuIyRGIaIleIeUyTrp 80
Db 252 CTCTCTCAAAAGCGCTTGAACAAGAACTCTTATGACTTTGAGAGATCTGTGCTGTGG 311
Qy 81 ThrLeuGIuPProGIuIyArGIaRGIaIeUyAlaIeUyAlaSnGIuAlaThrIySArGIyThr 100
Db 312 ACCCTGATCTTCGCGAGCTGATGCACTTCTTCCATTAACACTTCCAAAGATTTGACT 371
Qy 101 SerSerAenGIuValIeUyMetGIuValaIaCySThrArGIThrSerThrGIuIeUyHIs 120
Db 372 TCGAGCAACTGGGTTCTTCATGAAATGCTTGCACAGAGGCTTCAATGAGATTATTCATG 431
Qy 121 AlaRGIaIaIyThIsaIaArGIyRlySLeuSerIeUyGIuIyPValaIaIhIsHIs 140
Db 432 GTAGGCGAGCGCTATATAGCTGTGTATTAAGAAATCTTGAAGAAACATCGCATATCAC 491
Qy 141 ThrThrGIyAspPheArGIySLeuIySValIaSerIeUyAlaThrSerThrArGIyGIuGIy 160
Db 492 ACTACGCGGATTTCCGCAAGCTGCTTTCCTCTGCAAGTACCTTTCGATATGAGGG 551
Qy 161 AspGIuValaIaenMetThrLeuAlaIySgInuAlaIySLeuValIhIsgIuIySIlElys 180
Db 552 CTTGAGGTGAACATGACATGGCGAGATCAAGAGCTTAAGATCTCATGATGAGAAATTCAC 611
Qy 181 AspLyshIstYrAenAepGlValaIaIeArGIleuSerThrAgsSerIySAlaIeIn 200
Db 612 GAGAAAGCTTACATATGATAGTATGATCATCAGAAATTTGTTATCAAGAAAGTAAAGCTCAG 671
Qy 201 ILeaSnAlaThrPheaSnArGIyRGIhAspAspHIsGIyGluGluIleuIySerIeu 220
Db 672 CTTAATGCAACCTCAATTAATCTTCAACATGAAGTTTGGAAATGCCATCAACAGATCTG 731
Qy 221 GluGIuGIuIyAspAspAspIySPhelEuAlaIeuleuArGIserThriIeGIuCySleu 240
Db 732 ---AAGGCTGATCCAAATGATGAAATTTCTGAACCTCGTGCATCAACAAATTAAGTCTTG 788
Qy 241 ThrArpProGIuIeUyRrPheValaIaSpValIeUySeraIaIleAnIySThrGIyThr 260
Db 789 ACTTATCCAGAGAGATCTTTGAGAGTACTGCGCTGCGCATCAACAGCTGAGAAC 848
Qy 261 AspGIuGIuIyAlaIeUyThrArGIleValIThrThArGIaGIuIleAepIySValIle 280
Db 849 GATGAATGGGCTCTTTCAGAGATGACACACTCGTGGAGAGTTGACATGACGCGGATC 908
Qy 281 GIyGIuGIuIyRGIuArGIaRGIaSenSerIeProIeUySAlaIleThrIySAspThr 300
Db 909 AAGAAACAGTACCAAAAGAAAGAAAGAGTTCCTCTGATCGGCTATTCGAATGACAC 968
Qy 301 ArgGIyAspTYrGIuIySmetIeUyAlaIaIeUleuGIyGluAspAspAla 317
Db 969 TCTGGGAGCTTACAAAGAAAGATGCTTCTGCTTGGTTGAGACATGAGATGCT 1019
```

## RESULT 10

```
US-10-425-114-10772
; Sequence 10772, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425, 114
```

```
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ. ID NOS: 73128
/ SEQ. ID NO 10772
/ LENGTH: 1145
/ TYPE: DNA
/ ORGANISM: Glycine max
/ FEATURE:
/ OTHER INFORMATION: Clone ID: 700970311_FLI
US-10-425-114-10772

Alignment Scores:
Pred. No.: 3,83e-111 Length: 1145
Score: 1112.50 Matches: 225
Percent Similarity: 83.70% Conservative: 42
Best Local Similarity: 70.53% Mismatches: 49
Query Match: 69.06% Indels: 5
DB: 18 Gaps: 1

US-10-690-564-2 (1-317) x US-10-425-114-10772 (1-1145)

QY 1 MetAlaThrLeuIysValIserAspSerValProAlaProSerAspAspAlaGluGlnLeu 20
DB 54 ATGGCGACCTTAAGGTTCTCTACAGCTCTTCCCGCTTGACAGATGATGTGGACAGCTC 113
QY 21 ArgThrAlaPheGluGlyTrpGlyThrAsnGluAspLeuIleIleSerIleLeuAlaHis 40
DB 114 CGAAAGCCCTTCAGGTTGGGGAACCTAACGAGAGCTGATTTGATCGANTCTGGCTCAT 173
QY 41 ArgSerAlaGluGlnArgLysValIleArgGlnAlaIleThrHisGluThrTyrglyLysAsp 60
DB 174 AGGAATCTGCTCAGAGGAGGAGCTTAATCAGAAACTTATCCCAAGCTTATGAGAAAGAT 233
QY 61 LeuLeuIysThrLeuAspLysGluLeuSerAsnAspPheGluArgAlaIleLeuLeuTrp 80
DB 234 CTCTCTCAAGGCTTGACCAAGAACTCACGATGATTTTGAAGGCTGGTTCATCTTTGG 293
QY 81 ThrLeuGluProGlyIguArgAspAlaLeuLeuAlaAsnGluAlaThrLysArgTrpThr 100
DB 294 ACCTTGATCTCTGCTGACGTGATGATTTTGGCAACGAGCGACTAAATAATGAGCT 353
QY 101 SerSerAsnGlnValLeuMetGluValAlaCysThrArgThrSerThrGlnLeuLeuHis 120
DB 354 TCAGAGCAATCAGGTTCTGGTGGAAATAGCTGCATGAGGCTCTGTAACAATGTTGGCT 413
QY 121 AlaArgGlnAlaIleThrHisAlaArgTyrglyLysSerLeuGlnGluAspValAlaHisHis 140
DB 414 GCGAGGAAGCTTAACCATGTTCTTTATAGAGATCTCTGGAGAGATGTTGCTCATCAC 473
QY 141 ThrThrGlyAspPheArgLysLeuLeuValSerLeuValThrSerTyrglyArgTyrgly 160
DB 474 ACAACGAGAGACTTCCGTAAGCTCATACTACTCTGGTTAGTCTTATGCAATGAGGA 533
QY 161 ArgGluValAsnMetThrLeuAlaLysGlnGluAlaLysLeuValHisGluLysIleLys 180
DB 534 GATGAGGTCAACTTGACCTT-GCAAAAACCTGACCAAAAATTCGTCATGAGAGATTTC 592
QY 181 AspLysHisThrAsnAspGluAspValIleArgIleLeuSerThrArgSerLysAlaGln 200
DB 593 AACAGGCTTAATATGATGAGCACTTCATCAGATTTTGGCCACAGAGACAGACAGAC 652
QY 201 IleAsnAlaThrPhe-Asn-ArgTyrglnAspAspHisGlyGluGlnIleLeuLysSer 220
DB 653 ATTAAGCTACTTTTGAATCCACTACAAAGATGATTTGGCAAAATATATCAACAGGA-C 711
QY 220 euGluGluGlyAspAspAspAspLysPheLeuAlaLeuLeuArgSerThrIleGlnCysL 240
DB 712 TG---AAGGCTGATCCAAAGACAGAGTTCCTGCTTTACTTAAGAGCAACTGCTAAAGTCT 768
QY 240 euThrArgProGluLeuTyrglyPheValAspValLeuArgSerAlaIleAsnLysThrGly 260
DB 769 TGATCCGTCCCGAAGATCTCGAAGAGGTGTGTGGCTTGGGATCAACAGCGAGGA 828
QY 260 hrAspGluGlyAlaLeuThrArgIleValThrThrArgAlaGluIleAspLeuLysValI 280

RESULT 11
US-10-424-599-70457
/ Sequence 70457, Application US/10424599
/ Publication No. US20040031072A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa Thomas J
/ APPLICANT: Kovalic David K
/ APPLICANT: Zhou Yihua
/ APPLICANT: Cao Yongwei
/ TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
/ TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
/ FILE REFERENCE: 38-21 (53223)B
/ CURRENT APPLICATION NUMBER: US/10/424,599
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ. ID NOS: 285684
/ SEQ. ID NO 70457
/ LENGTH: 1321
/ TYPE: DNA
/ ORGANISM: Glycine max
/ FEATURE:
/ OTHER INFORMATION: Clone ID: PAT_MRT3847_34633C.1
US-10-424-599-70457

Alignment Scores:
Pred. No.: 7.41e-104 Length: 1321
Score: 1046.50 Matches: 213
Percent Similarity: 81.56% Conservative: 48
Best Local Similarity: 66.56% Mismatches: 55
Query Match: 64.96% Indels: 4
DB: 18 Gaps: 1

US-10-690-564-2 (1-317) x US-10-424-599-70457 (1-1321)

QY 1 MetAlaThrLeuIysValIserAspSerValProAlaProSerAspAspAlaGluGlnLeu 20
DB 90 ATGGCGACCTTAAGGTTCTCTACAGCTCTTCCCTTCCATTGGAAGACAGTACGATTTG 149
QY 21 ArgThrAlaPheGluGlyTrpGlyThrAsnGluAspLeuIleIleSerIleLeuAlaHis 40
DB 150 CGAAAGCCCTTTCAAGGGTGGGGAACGAATGAGATGATGATATCAATCTGGGCTCAC 209
QY 41 ArgSerAlaGluGlnArgLysValIleArgGlnAlaIleThrHisGluThrTyrglyLysAsp 60
DB 210 AGAAATGCGGCTCAGGCTTAAGTTCGTAAGCTTATGATCAACCAATGAGAGAGAT 269
QY 61 LeuLeuIysThrLeuAspLysGluLeuSerAsnAspPheGluArgAlaIleLeuLeuTrp 80
DB 270 CTCTTTAAAGATTATGACAAAGAACTTTCATGACTTTGACCGGGCTGTGTGGTTGG 329
QY 81 ThrLeuGluProGlyIguArgAspAlaLeuLeuAlaAsnGluAlaThrLysArgTrpThr 100
DB 330 ACATTTGATCCGCTGAGGCTGATCCCTTTTATGATTAATGAAGCCACTAAGATGTTGACT 389
QY 101 SerSerAsnGlnValLeuMetGluValAlaCysThrArgThrSerThrGlnLeuLeuHis 120
DB 390 TCAACCAATTTGGGTCAATTTTGGAAATGCTTCATGATCTTCACTTGATCTCTTAAG 449
QY 121 AlaArgGlnAlaIleThrHisAlaArgTyrglyLysSerLeuGlnGluAspValAlaHisHis 140
DB 450 GDAAGCAAGCAATATCAACCCGCTTCAAAAAGTCCCTTGAAGAGACCTTCCATTCAT 509
QY 141 ThrThrGlyAspPheArgLysLeuLeuValSerLeuValThrSerTyrglyArgTyrgly 160
```

```
Db 510 ACTAAGGGGACATCCCGACGCTATTGTCCTCTTGAGCATATCCGTTATGAGGG 569
Qy 161 Asp-GluValAspMetThrLeuAlaLysGlnGluAlaLysLeuValHisGluLysIleLy 180
Db 570 GATTTGAGGTGAACCTGATCTGGCAAAATCTGAGGGCTAACTGCTTCCACGAAAGATTGC 629
Qy 180 GAAPLVHIEYrAspAspGluAspValIleArgIleLeuSerThrArgSerLysAlaG1 200
Db 630 AGAGAGGCGCTACATGATGAGAGCTGATCGATTTTGTCAACAGAGTAAAGACA 669
Qy 200 nileAsnAlaThrPheAsnArgTyrglnAspAspHisGlyGluGluIleLeuLysSerLe 220
Db 690 GTTAAAGCAACACTAATGATCAGTACATGATGAGTTGGGAATGCCATAACAAGATCT 749
Qy 220 uGluGlu--GlyAspAspAspAspLysPheLeuAla-LeuLeuArgSerThr-IleGln 238
Db 750 GAAAAATTGAACCTCAAAAAGAAATGTAATCTTGCAATTTATTGAGGGCAAGCATTAAG 809
Qy 239 CysLeuThrArgProGluLeuLeuTyrglnPheValAspValLeuArgSerAlaIleAsnLysThr 258
Db 810 TGCTGCTTAACCTTGAAATATTTCGAAAAGTCTTGAGGCTGGCTATTAACAAGCTA 869
Qy 259 GlyThrAspGluGluAlaLeuThrArgIleValThrThrArgAlaGluIleAspLeuLys 278
Db 870 GGAATCATGATAAGGGGCGCTTACTAGAGTGTGACACACAGAGCTGAGTTGATTGTCAG 929
Qy 279 ValIleGlyGluGluTyrglnArgArgAsnSerIleProLeuGluGluValaIleThrLys 298
Db 930 CGAATGCTGAGGAATCTCAAAAGAAACAGCATTCCTCTGAGCCGCAATAGCCAGT 969
Qy 299 AspThrArgGlyLysArgTyrgluLysMetLeuValAlaLeuLeuGlyGluAspArg 316
Db 990 GACACTTCTGAGACATACAGACATTCCTCTGACATGTTGACATGAGGAT 1043
```

## RESULT 12

```
US-09-770-149-655
; Sequence 655. Application US/09770149
; Patent No. US20020059663A1
; GENERAL INFORMATION:
; APPLICANT: Gorlach, Jörn
; APPLICANT: An, Yong-Qiang
; APPLICANT: Hamilton, Carol M.
; APPLICANT: Price, Jennifer L.
; APPLICANT: Raines, Tracy M.
; APPLICANT: Yu, Yang
; APPLICANT: Rameaka, Joshua G.
; APPLICANT: Page, Amy
; APPLICANT: Matthew, Abraham V.
; APPLICANT: Ledford, Brooke L.
; APPLICANT: Woessner, Jeffrey P.
; APPLICANT: Haas, William David
; APPLICANT: Kricker, Carlos A.
; APPLICANT: Slader, Ted
; APPLICANT: Davis, Keith R.
; APPLICANT: Allen, Keith
; APPLICANT: Hoffman, Neil
; APPLICANT: Hurban, Patrick
; TITLE OF INVENTION: Expressed Sequences of Arabidopsis
; TITLE OF INVENTION: thaliana
; FILE REFERENCE: 2024 (PARA-013PRV)
; CURRENT APPLICATION NUMBER: US/09/770,149
; PRIOR FILING DATE: 2001-01-26
; PRIOR FILING DATE: 2000-01-27
; NUMBER OF SEQ ID NOS: 999
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 655
; LENGTH: 635
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-770-149-655
```

## Alignment Scores:

```
Pred. No.: 3,9e-104 Length: 635
Score: 1045.00 Matches: 205
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 64.87% Indels: 0
DB: 9 Gaps: 0
```

US-10-690-564-2 (1-317) x US-09-770-149-655 (1-635)

```
Qy 1 MetAlaThrLeuLysValSerAspSerValProAlaProSerAspAspAlaGlnGlnLeu 20
Db 19 ATGGCAGACTCTTAAGTTTCTGATTTCTGTTCTTCCTGCTCTTCTGATGATGCGCAATTG 78
Qy 21 ArgThrAlaPheGluGluTyrglyThrArgGluAspLeuIleIleSerIleLeuAlaHis 40
Db 79 AGAACCGCTTTGAAAGATGGGGGTACGAACGAGGACTTGATCATATCAATCTTGCGCTCAC 138
Qy 41 ArgSerAlaGlnGlnArgLysValIleArgGlnAlaThrHisGluThrTyrglyGluAsp 60
Db 139 AGAAGTGTGAAACAGAGAAAGTCATCAGGCAAGCATACACGAAACCTACGCGCAAGAC 198
Qy 61 LeuLeuLysThrLeuAspLysGluLeuSerAspAspPheGluLysArgAlaIleLeuLeuTrp 80
Db 199 CTTCTCAAGACTCTTACCAAGAGCTCTTACGATTTGAGAGAGCTATCTTGTGTGG 258
Qy 81 ThrLeuGluProGlyGluArgAspAlaLeuLeuAlaAsnGluAlaThrLysArgTrpThr 100
Db 259 ACTCTTGAAACCGGTGAGGTGATGTTATTGGCTTAATGAGCTCAAAAAGATGGAGCT 318
Qy 101 SerSerAsnGlnValLeuMetGluValAlaCysThrArgThrSerThrGlnLeuLeuHis 120
Db 319 TCAAGCAACCAAGTTCTTATGAAAGTCTTGCACAAAGACATCAACGAGCTGCTTAC 378
Qy 121 AlaArgGlnAlaTyrglnAlaArgTyrglyLysSerLeuGluGluAspValAlaHisHis 140
Db 379 GCTAGGCAAGCTTACCATGCTGCTGCTACAGAAAGTCTCTTGAAGAGAGAGCTTGCTCAC 438
Qy 141 ThrThrGlyAspPheArgLysLeuLeuValSerLeuValThrSerTyrglyTyrgly 160
Db 439 ACTACCGGTGACTTCAAGAAAGCTTTGGTTTCTCTTGAACCTCATACAGAGTCAAGGA 498
Qy 161 AspGluValAspMetThrLeuAlaLysGlnGluAlaLysLeuValHisGluLysIleLys 180
Db 499 GATGAAGTGAACATGATGCTTGAAGCAAGAGCTTACGTGTCATAGAAATCAAG 558
Qy 181 AspLysHisTyrglnAspGluAspValIleArgIleLeuSerThrArgSerLysAlaGln 200
Db 559 GACAGCACTACATGATGAGATGTTATTGAATCTTGTCCACAGAAAGCAAGAGCTCAC 618
Qy 201 IleAsnAlaThrPhe 205
Db 619 ATCAATGCTACTTTT 633
```

## RESULT 13

```
US-10-767-701-13525
; Sequence 13525. Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5353)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 13525
; LENGTH: 1385
; TYPE: DNA
; ORGANISM: Sorghum bicolor
```



```

FEATURE:
OTHER INFORMATION: Clone ID: SORBI-28MAY03-CLUS5887_1
US-10-767-701-13525

Alignment Scores:
Pred. No.: 5,21e-100 Length: 1385
Score: 1011.50 Matches: 192
Percent Similarity: 77.78% Conservative: 53
Best Local Similarity: 60.95% Mismatches: 69
Query Match: 62.79% Indels: 1
DB: 19 Gaps: 1

US-10-690-564-2 (1-317) x US-10-767-701-13525 (1-1385)

QY 1 MetAlaThrLeuIysValSerAspSerValProAlaProSerAspAlaGluGlnLeu 20
DB 58 ATGGCGAGCTCAAGGCTCCAGCGCATCTCCCTCCGTCAGATGATCGGAGCAGCTC 117
QY 21 ArgThrAlaPheGluGlyTrpGlyThrAsnGluAspLeuIleIleSerIleLeuAlaHis 40
DB 118 CGCAGAGCATTCACAGGCTGGGCGACAGAGCGCTCATCTCTCCATCTCGGCCAC 177
QY 41 ArgSerAlaGluGlnArgValIleArgGlnAlaTrpHisGluThrTyrglyGluAsp 60
DB 178 CGCGACCCCGCGAGCGCCGCGCATCCGCGCGCTACCGCCAGCGCCAGCGGAGAG 237
QY 61 LeuLeuIysThrLeuAspIysGluLeuSerAspPheGluArgAlaIleLeuLeuTrp 80
DB 238 CTCTCCGCTCCATCAGCAGAGATCTCCGCGCACTTCAGAGAGCGCGTGAATCTGTGG 297
QY 81 ThrLeuGluProGlyGluArgAspAlaLeuLeuAlaAsnGluAlaThrIysArgTrpThr 100
DB 298 ACGCTGGACCCAGAGCGGCGCGACGCGGTCTGGCCAGAGCGGCGGAGAGTGCGAG 357
QY 101 SerSerAsnGlnValLeuMetGluValAlaCysThrArgThrSerThrGlnLeuLeuHis 120
DB 358 CCAAGGAACCGCGTCTGCTCGAGATCGCTGACGCGCACTTCGCGCGAGGTCTTCGCG 417
QY 121 AlaArgGlnAlaTrpHisAlaArgTrpIysSerLeuGlnGluAspValAlaHisHis 140
DB 418 GCCAGGCGAGCGTATCCAGAGCGGTTCAGCGCTCGCTTAGAGAGCATCGCGCCAC 477
QY 141 ThrThrGlyAspPheArgIysLeuLeuValSerLeuValThrSerTrpArgTrpGlyGlu 160
DB 478 GTCACTGGGACCTTCGCAAGCTACTGCTGCCACTTGAATGATCATCCGCTATGATGA 537
QY 161 AspGluValaIasMetThrLeuAlaIysGlnGluAlaIysLeuValHisGluIysIleIys 180
DB 538 CCAGAGGTCAACACAGAGTGGCACACTCAGAGGCCAACTACTCCATGGAAGATCCAT 597
QY 181 AspIysHisIysTrpAsnAspGluAspValIleArgIleuSerThrArgSerIysAlaGln 200
DB 598 CACAGGCTTACATGATGATGATGATCATCAAGATCTTCACACTAGAGCAACAACTCAG 657
QY 201 IleAsnAlaThrPheAsnArgTrpGlnAspAspHisGlyGluGlnIleLeuIysSerLeu 220
DB 658 CTGCTCTCATCATTCATCATTCATCAATGATGATTCGCAACCGAATCAACAAAGATCTG 717
QY 221 GluGluGluIysAspAspAspIysPheLeuAlaLeuLeuArgSerThrIleGlnCysLeu 240
DB 718 --AAGGCTACCCCAAGATGAGTACTCAAAACATTCGGGGCAATCATCCGCTGCTTC 774
QY 241 ThrArgProGluLeuTrpPheValAspValLeuArgSerAlaIleAsnIysThrGlyTrp 260
DB 775 AGTTGCCCCCAACAGTACTTTGAGAAGTTCGCCAGGCGCATAGCGGGGCTTAGGACA 834
QY 261 AspGluGluValaLeuThrArgIleValThrThrArgAlaGluIleAspLeuIysValIle 280
DB 835 GACGAGAAACCGCTGACAGGCTCATATCACCGCGGCTGAGGTGAGCATGAAACGATTT 894
QY 281 GlyGluGluIysTrpGlnArgAsnSerIleProLeuGlnIysAlaIleThrIysAspThr 300
DB 895 AAGAGGCGTACCAAGAGAAACAGCGTCCCTGAGCGCGCGTCTGGGGACACC 954

```

```

QY 301 ArgGlyAspTrpGluIysMetLeuValAlaLeuLeuGlyGluAsp 315
DB 955 TCCGGGCACTACGAGAGCATGCTCTCCGCGCTCTTGGGCAAGAG 999

RESULT 14
US-10-425-114-21986
Sequence 21986, Application US/10425114
Publication No. US20040034888A1
GENERAL INFORMATION:
APPLICANT: Liu, Jindong
APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E.
APPLICANT: Tabaka, Jack E.
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(5313)B
CURRENT APPLICATION NUMBER: US/10/425,114
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 21986
LENGTH: 1235
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: LIB3356-011-F3_FLI
US-10-425-114-21986

```

```

Alignment Scores:
Pred. No.: 7.34e-100 Length: 1235
Score: 1009.50 Matches: 191
Percent Similarity: 77.46% Conservative: 53
Best Local Similarity: 60.63% Mismatches: 70
Query Match: 62.66% Indels: 1
DB: 18 Gaps: 1

US-10-690-564-2 (1-317) x US-10-425-114-21986 (1-1235)

QY 1 MetAlaThrLeuIysValSerAspSerValProAlaProSerAspAlaGluGlnLeu 20
DB 78 ATGGCAACCTCAAGGCTCCAGCGCATCTCTCCGTCGCGAGTACTGCCACGAGCTG 137
QY 21 ArgThrAlaPheGluGlyTrpGlyThrAsnGluAspLeuIleIleSerIleLeuAlaHis 40
DB 138 CGAAGCATTCAGAGGCTGGGACAGAGCGGCTCATATCTCCATCTCCAGGCCAC 197
QY 41 ArgSerAlaGluGlnArgValIleArgGlnAlaTrpHisGluThrTyrglyGluAsp 60
DB 198 CGGAGCGCGCGGAGCGCGCCCATCCGCGGCTTACGCCGAGGCGCTACGCGGAGAG 257
QY 61 LeuLeuIysThrLeuAspIysGluLeuSerAspPheGluArgAlaIleLeuLeuTrp 80
DB 258 CTCTCCGCTCATCAACGAGCATCTCCGCGCACTTCGAGAGGCGCGTGAATCTGTGG 317
QY 81 ThrLeuGluProGlyGluArgAspAlaLeuLeuAlaAsnGluAlaThrIysArgTrpThr 100
DB 318 ACCTTCCAGCCAGCGGAGCGGAGGCTGCTGGCCCAAGAGCGCGCGGAGAGTGAAG 377
QY 101 SerSerAsnGlnValLeuMetGluValAlaCysThrArgTrpSerThrGlnLeuLeuHis 120
DB 378 CCGCGGAAACGAGTGTCTCGAGATCGCTCCAGCGGCACTCCGCGGCAATCTTCGCG 437
QY 121 AlaArgGlnAlaTrpHisAlaArgTrpIysIysSerLeuGlnGluAspValAlaHis 140
DB 438 ACCAGGCAAGCTTATACAGAGCGGTTCAGAGCGCTCGCTTGAAGAAACATCTCGCGCAC 497
QY 141 ThrThrGlyAspPheArgIysLeuLeuValSerLeuValThrSerTrpArgTrpGlyGlu 160
DB 498 GTCACTGGGGACTTCGCAAGGCTATGTGGCCACTTGATGATCTTACCGCTATGATGA 557
QY 161 AspGluValaIasMetThrLeuAlaIysGlnGluAlaIysLeuValHisGluIysIleIys 180

```

```

Db      558 CCAGAGCTCAACAGAGCTTGCACACTCAGAACCCAACTCTTCATGAGAAATCAT 617
Qy      181 AsplyshiSTYraenapgluapvalllearglleuserthrargserlysalglu 200
Db      618 CACAAAGCTTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 677
Qy      201 lIeasnlathrPheasnaargTyrGlnaspsphiegllygluullleuylserleu 220
Db      678 CTGATTCTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 737
Qy      221 GluylugluYraapapapapapapapapapapapapapapapapapapapapapap 240
Db      738 ---AAGCTGACCCCGACGAGAGATCTCAGAACGCTCGCGGAGATCTCGGCTTC 794
Qy      241 ThrArgProgluLeuYrPheValAspValleuArgSerThrlleasnlYthrGlyThr 260
Db      795 AGTTGCCCTGACAGATCTTTGAGAGAGTGGCTAGGAGGCTATAGCGGGCTAGGCACA 854
Qy      261 AspgluGluYalaleuthrArglleValThrThrArgAlagluileaspleuysValile 280
Db      855 GATGAGAACTCCCTGACGAGGTCATCATCACCCGCTGAGGTGAGACTCAAACTGATT 914
Qy      281 GlygluGluYrGlnArgAsnSerlleProleuGluYalaleuthrlyAspThr 300
Db      915 AAGAGGCGCTACCGAAGAGAACAGAGCGCTCGCTGAGCGGCGCTCGCGGAGACAC 974
Qy      301 ArgGlyAspTyrGluYrMetLeuValAlaleuLeuGlyluasp 315
Db      975 TCCGGCGACTACGAGAGCATGCTCTCGCGCTCGGGGCGAGAG 1019

RESULT 15
US-10-425-114-26803
; Sequence 26803, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OR INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 26803
; LENGTH: 1285
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB4605-011-H2_FLI
US-10-425-114-26803

Alignment Scores:
Pred. No.: 7,75e-100 Length: 1285
Score: 1009.50 Matches: 191
Percent Similarity: 77.46% Conservative: 53
Best Local Similarity: 60.63% Mismatches: 70
Query Match: 62.66% Indels: 1
DB: 18 Gaps: 1

US-10-690-564-2 (1-317) x US-10-425-114-26803 (1-1285)
Qy      1 MetAlathreulysValserAspSerValProAlaProSerAspAspAlaGluGlnLeu 20
Db      144 ATGGCAACGCTCAAGGCTCCAGCCACCTGCTCCCTCCGCGATGACTCGACCAAGCTG 203
Qy      21 ArgThrAlaPheGluGlyTyrGlyThrAsnGluAspLeuIleIleSerIleLeuAlaHis 40
Db      204 CGCAAGCATTTCCAGGCTGGGGCAGACGAGCGCTCATCATCTTCATCTTCGCGCAC 263

```

```

Qy      41 ArgSerAlaGluGlnArglyValIleArgGlnAlaTyrHisgluThrTyrGlyGluAsp 60
Db      264 CGGAGCGCCCGGAGCGCCCGCCATCCGCGCTTACCCGAGGCTTCAGCGGAGAG 323
Qy      61 LeuLeuYrThrleuAspYrGluLeuSerAspPheGluArgAlaIleLeuLeuTyr 80
Db      324 CTCTCGCTCCATTCACGAGAGATCTCCGGGACTTCGAGAGGGCCGATCTCTGG 383
Qy      81 ThrLeuGluProGlyGluArgAspAlaLeuLeuAlaAsnGluAlaThrlySarGlyPThr 100
Db      384 ACCGTGACCCAGCGAGCGCGACGCGGTCTGCGCAAGAGCGGCGCGGAAGTGGAG 443
Qy      101 SerSerAsnGlnValLeuMetGluValAlaCysThrArgThrSerThrGlnLeuHis 120
Db      444 CCGCGGAAACGTGTGCTCGTGAAGATCGCTGACGCGCACCTCCGCGCAGATCTTCGCG 503
Qy      121 AlaArgGlnAlaThrHisAlaArgTyrlyValSerleuGluGluAspValAlaHisHis 140
Db      504 ACCAGGAGGCTTATACAGAGCGGTTCAGCGCTCGCTGAGAGAGCATCGCTGGCAC 563
Qy      141 ThrThrGlyAspPheArglyLeuLeuValSerleuValThrSerTyrArgTyrGlyGly 160
Db      564 GTCACTGGGAGCTTCGCAAGATGATGTCCTGTAAGTACTTACCGCTATGATGA 623
Qy      161 AspgluValAsnMetThrleuAlalyGlnGluAlalyLeuValHisgluYrlyleYs 180
Db      624 CCAGAGCTCAACCAAGGTTGGCACCTCAGAACCCAACTCATGATGAGAAATCCAT 683
Qy      181 AsplyshiSTYraenapgluapvalllearglleuserthrargserlysalglu 200
Db      684 CACAAAGCTTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 743
Qy      201 lIeasnlathrPheasnaargTyrGlnaspsphiegllygluullleuylserleu 220
Db      744 CTGATTGCTACATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 803
Qy      221 GluylugluYraapapapapapapapapapapapapapapapapapapapapapap 240
Db      804 ---AAGCTGACCCCGACGAGAGATCTCAGAACCTCGCGGCGATCTCCGCTCTTC 860
Qy      241 ThrArgProgluLeuYrPheValAspValleuArgSerThrlleasnlYthrGlyThr 260
Db      861 AGTTGCCCTGACAGATCTTTGAGAGAGTGGCTAGGAGCGCATAGCGGGCTAGGCACA 920
Qy      261 AspgluGlyAlaleuthrArglleValThrThrArgAlagluileaspleuysValile 280
Db      921 GATGAGAACTCCCTGACGAGGTCATCATCACCCGCTGAGGTGAGACTCAAACTGATT 980
Qy      281 GlygluGluYrGlnArgAsnSerlleProleuGluYalaleuthrlyAspThr 300
Db      981 AAGAGGCGCTACCGAAGAGAAACAGCGTGCCTGAGCGGCGCTCGCGGAGACACC 1040
Qy      301 ArgGlyAspTyrGluYrMetLeuValAlaleuLeuGlyluasp 315
Db      1041 TCCGGCGACTACGAGAGCATGCTCTCGCGCTCGGGGCGAGAG 1085

```

Search completed: August 22, 2005, 14:49:41  
Job time : 762 secs

**This Page Blank (uspto)**